BAYESIAN HIERARCHICAL, SEMIPARAMETRIC, AND NONPARAMETRIC METHODS FOR INTERNATIONAL NEW PRODUCT DIFFUSION

A Dissertation

by

BRIAN MATTHEW HARTMAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2010

Major Subject: Statistics

BAYESIAN HIERARCHICAL, SEMIPARAMETRIC, AND NONPARAMETRIC METHODS FOR INTERNATIONAL NEW PRODUCT DIFFUSION

A Dissertation

by

BRIAN MATTHEW HARTMAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Bani K. Mallick
Committee Members,	David B. Dahl
	Larry G. Gresham
	Jeffrey D. Hart
	Debabrata Talukdar
Head of Department,	Simon J. Sheather

August 2010

Major Subject: Statistics

ABSTRACT

Bayesian Hierarchical, Semiparametric, and Nonparametric Methods for International New Product Diffusion. (August 2010)

Brian Matthew Hartman, B.S., Brigham Young University

Chair of Advisory Committee: Dr. Bani K. Mallick

Global marketing managers are keenly interested in being able to predict the sales of their new products. Understanding how a product is adopted over time allows the managers to optimally allocate their resources. With the world becoming ever more global, there are strong and complex interactions between the countries in the world. My work explores how to describe the relationship between those countries and determines the best way to leverage that information to improve the sales predictions.

In Chapter II, I describe how diffusion speed has changed over time. The most recent major study on this topic, by Christophe Van den Bulte, investigated new product diffusions in the United States. Van den Bulte notes that a similar study is needed in the international context, especially in developing countries. Additionally, his model contains the implicit assumption that the diffusion speed parameter is constant throughout the life of a product. I model the time component as a nonparametric function, allowing the speed parameter the flexibility to change over time. I find that early in the product's life, the speed parameter is higher than expected. Additionally, as the Internet has grown in popularity, the speed parameter has increased.

In Chapter III, I examine whether the interactions can be described through a reference hierarchy in addition to the cross-country word-of-mouth effects already in the literature. I also expand the word-of-mouth effect by relating the magnitude of the effect to the distance between the two countries. The current literature only

applies that effect equally to the n closest countries (forming a neighbor set). This also leads to an analysis of how to best measure the distance between two countries. I compare four possible distance measures: distance between the population centroids, trade flow, tourism flow, and cultural similarity. Including the reference hierarchy improves the predictions by 30% over the current best model.

Finally, in Chapter IV, I look more closely at the Bass Diffusion Model. It is prominently used in the marketing literature and is the base of my analysis in Chapter III. All of the current formulations include the implicit assumption that all the regression parameters are equal for each country. One dollar increase in GDP should have more of an effect in a poor country than in a rich country. A Dirichlet process prior enables me to cluster the countries by their regression coefficients. Incorporating the distance measures can improve the predictions by 35% in some cases.

To Jenn, my constant support and eternal companion

ACKNOWLEDGMENTS

First I would like to thank my advisor, Bani Mallick, for his constant support and kindness. He was always willing to overlook the simple errors I would make and kindly put me back on track. He taught me not only about statistics, but also how to be an academic. His lessons will help me throughout my career.

Also, thanks to Debu Talukdar my always willing and ever-present guide to the wonderful world of marketing. He was indispensable in my education and this work. Without him, I would have never known how exciting international new product diffusion really is.

I am grateful for David Dahl who introduced me to the amazing computational advantages of Scala and the vibrant world of Bayesian Nonparametrics.

Jeffrey Hart and Larry Gresham were always willing to offer support and guidance.

I am also grateful to Craig Meisner (World Bank) and Azucena Pernia (United Nations World Tourism Organization) for their help in providing much of the data used in this study.

Finally, I appreciate the support of the National Science foundation CMG research grant DMS-0724704.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

A. Problem

Global marketing managers are deeply concerned with predicting how well their products will sell. As the world has become more global [1], it is more important for managers to use relevant information from other countries to improve their predictions. I discuss various aspects of international new product diffusion which are related to the flat world in which we currently live. I examine the best way to describe the cross-country influence, how to measure the distance between two countries, how the diffusion speed changes throughout the life of the product and with the proliferation of the internet, and how to best cluster the countries into groups to improve the predictions. The remainder of this chapter describes the data set of which I use various subsets of to solve each problem.

B. Data

I collected relevant new product diffusion data for seven product categories across 31 countries. The product categories are microwave, fax machine, VCR, CD player, camcorder, home computer, and cellular phone.

Collecting data for international new product diffusion studies remains a challenging task and my own experience for this is no exception. While adoption (first purchase) data is ideal for estimating diffusion models, such data is very difficult to collect across a wide range of countries, especially for developing countries [2]. I was

---

The journal model is *IEEE Transactions on Automatic Control.*

able to obtain adoption data for the cell phone and personal computer categories. Conversely, for the CD player, fax machine, microwave, VCR, and cell phone categories, I use sales data. To reduce the impact of repeat purchases, I only use sales data for the first ten years of product life in a country. Under these conditions, the data series for each product begins and ends as described in Table I. Table II lists the

Table I. Product Date Ranges

Product	Start Date	End Date
Camcorder	1987	1996
CD Player	1985	1993
Cellular Phone	1981	2002
Fax Machine	1978	1991
Microwave	1975	1991
Personal Computer	1981	2004
VCR	1977	1987

31 countries that I use in my study. The list consists of most of the major developed and developing countries, and together account for about 80% of the world economic output and 60% of the world population. Taken together, my study has 217 (7x31) product and country pairs with a broad representation in terms of key developed and developing countries. In the context of international diffusion studies, the scope of my data provide a substantial empirical basis for investigation. For instance, [3] notes that a substantial data basis in this context should have a sample size of more than ten countries or ten products.

My data was obtained from several international organizations including the International Monetary Fund (IMF), International Telecommunications Union (ITU), the United Nations (UN), the World Bank and the World Tourism Organization (WTO). Product sales data are obtained from the databases of the World Bank, ITU, and publications by Euromonitor (European and International Marketing Data

Table II. Countries in My Sample

Country	% Pop.	% GNI	Country	% Pop.	% GNI
Argentina	0.6	0.96	Italy	0.91	3.06
Australia	0.31	1.19	Malaysia	0.39	0.47
Austria	0.13	0.49	Mexico	1.6	1.96
Belgium	0.16	0.62	Netherlands	0.25	1.01
Brazil	2.88	2.75	Norway	0.07	0.34
Canada	0.5	1.9	Philippines	1.29	0.83
Chile	0.25	0.32	Portugal	0.16	0.39
China	20.19	15.87	Singapore	0.07	0.22
Denmark	0.08	0.33	South Korea	0.75	1.91
Finland	0.08	0.3	Spain	0.67	2.05
France	0.94	3.5	Sweden	0.14	0.53
Germany	1.28	4.44	Switzerland	0.12	0.52
Greece	0.17	0.46	Thailand	0.99	0.98
Hong Kong	0.11	0.43	United Kingdom	0.93	3.65
India	16.94	6.73	United States	4.59	22.3
Ireland	0.06	0.25	**TOTAL**	**57.62**	**80.76**

and Statistics, various years) as well as various national government agencies. The UN and World Bank databases served as the source for various country-specific covariates. I have 12 time-invariant and 10 time-varying country-specific covariates, listed in Table III. By utilizing time-varying parameters I can investigate what factors influence the speed changes throughout the life of the product. In contrast to many of the existing new product diffusion studies, a unique feature and contribution of our study is assembling detailed information to measure the basis of bilateral interactions among countries analyzed. As noted in the model section, we use four different empirical measures as the basis for such bilateral interaction between countries: (1) the reciprocal of the distance between the population centroids for the two countries, (2) the tourism flow (in terms of number of tourists) between the two countries, (3) the trade flow (in constant dollar terms) between the two countries, and (4) the reciprocal of the cultural distance. Albuquerque et al. also used all but the tourism flow. Pop-

Table III. Possible Covariates

Time-invariant Covariates	**Time-varying Covariates**
Daily newspapers	Age dependency ratio
Ease of doing business index	Consumer price index
GINI index	Electric power consumption
Households with television	GDP per capita
Individualism Index	Household final consumption
International migration stock	Internet users
International tourism	Labor force participation rate, female
International voice traffic	Telephone mainlines
Pervasiveness of existing adopters	Unemployment, total
Population growth rate	Urban population
Price basket for residential fixed line	
Pump price for gasoline	
Uncertainty Avoidance Index	

ulation centroid data was obtained from the Socioeconomic Data and Applications Center at Columbia University. The population centroid is the geographical point which is nearest to all the people in the country, on average. We obtained a longitude and latitude for the centroid of each country and then calculated the distance between them using the Haversine formula as described in the appendix.

The bilateral tourism flow data was collected primarily from the database of the World Tourism Organization, but also from respective national tourism agencies. The bilateral trade flow data was collected from the general database of the United Nations Conference on Trade and Development (UNCTAD) and the Direction of Trade Statistics database of the International Monetary Fund. For our purpose, time-averaged values of annual tourism and trade flow levels are used. The annual averages of total tourism and trade flow levels among our entire sample of 31 countries are about 1 billion tourists and 8.3 trillion dollars respectively. As expected, there exists considerable variation across the 31 countries in terms of respective total tourism and trade flow levels. Also, for any given country, its levels of bilateral tourism and trade flows vary significantly across the other 30 countries. For instance, the coefficient of variations of China's bilateral tourism and trade flows with the other 30 countries are 4.6 and 9.7 respectively.

CHAPTER II

INVESTIGATING INTERNATIONAL NEW PRODUCT DIFFUSION SPEED: A SEMIPARAMETRIC APPROACH

A. Introduction

Global marketing managers are constantly concerned with how well their products will sell. There are many models in the literature which can be used to predict the future sales of a product. The majority of these models assume that the parameters associated with the diffusion speed are constant over time. I generalize a model currently used in the new product diffusion literature to allow the speed parameter to change over time. Using this model allows us to investigate how the speed has changed and what factors contribute to or describe the change.

Not surprisingly, there have been studies on the aforesaid issues, but are still limited to and almost exclusively set in the United States. [4] is the most recent major study which investigates new product diffusions in the United States. As Van den Bulte notes, an important research need is a similar study in an international context, especially developing countries. In addition to his observation, similar research has become more relevant in the context of accelerated globalization trends since the 1990s with the end of the cold war and the emergence of the internet [1]. My study works to fill that need.

My study covers four new product diffusions across 31 developed and developing nations from 1980-2004. My set of 31 countries accounts for about 80% of the global economic output and 60% of the global population. My study not only provides the needed and interesting international counterpart to the United States-focused study of [4] on change in new product diffusion speed, but also uses a novel methodological approach to analyze such changes in diffusion speed. While Van den Bulte's model contains the implicit assumption that the diffusion speed parameter is constant throughout the life of a product, I use semi-parametric regression to allow the speed parameter to change over the life of each product.

My analysis and the scope of my data enable us to gain insights into several important issues that are of interest to both marketing managers and researchers. They include: What key macro-environmental factors influence global new product diffusion speed? How does the speed of diffusion change over time for a given product as it diffuses across countries? To what extent is such acceleration due to changes in the levels of macro-environmental factors themselves versus an intrinsic change independent of the factors? These questions are currently very interesting as the past 25 years experienced major socio-economic events in the world with likely consequences on the propensity to adopt new products.

B. Data

As I am interested in the change in speed with the technological revolution of the 1990s, I only use the product information which covers that period. The four product categories are CD players, camcorders, home computers, and cellular phones.

C. Methodology

1. Model

I use the logistic diffusion model as my base model for the new product diffusion:

$$\frac{y(t)}{Y(t-1)} = \lambda \left[1 - \frac{Y(t-1)}{M(t)\alpha} \right] + \epsilon(t) \tag{2.1}$$

where $y(t)$ is the number of adopters in time t, $Y(t-1)$ is the number of cumulative adopters by time $t-1$, $M(t)$ is the population at time t, α is the proportion of the population which will eventually adopt the product(the adoption ceiling), λ is the parameter directly related to the diffusion speed, the main focus of this chapter, and $\epsilon(t)$ is the error term, $\epsilon(t) \sim N(0, \sigma^2)$. Previous work [4] examined 31 electrical household durables in the United States over a period of 74 years. His model for product n is:

$$\frac{y_n(t)}{Y_n(t-1)} = \lambda_n \left[1 - \frac{Y_n(t-1)}{M(t)\alpha_n}\right] + \sum_{k \in P_n} \beta_{kn}(X_{knt} - \bar{X}_{kn\cdot}) + \epsilon_n(t) \tag{2.2}$$

$$\lambda_n = \lambda_0 + \sum_{k \in P_n} \beta_k(\bar{X}_{kn\cdot} - \bar{X}_{k\cdot\cdot}) + \tau_n. \tag{2.3}$$

where β_k is the regression coefficient for covariate X_k and P_n is the set of product-specific covariates. To take into account my multiple countries and multiple products, I rewrite the model for country i and product n as:

$$\frac{y_{in}(t)}{Y_{in}(t-1)} = \lambda_{in} \left[1 - \frac{Y_{in}(t-1)}{M_i(t)\alpha_{in}}\right] + \epsilon_{in}(t) \tag{2.4}$$

$$\epsilon_{in}(t) \sim N(0, \theta_L). \tag{2.5}$$

Van den Bulte's model implicitly assumes that the speed parameter for a given product is constant throughout the life of the product. Throughout the life of a product, the speed parameter can change due to changes in the covariate values and other influences (current events, changes in advertising expenditures and competing products) exogenous to my model. By allowing the speed parameter to vary over the life of the product, I have more parameters than data points. My model becomes estimable through shrinkage [5], operationalized through a hierarchical structure:

$$\log \lambda_{in}(t) = f(t) + A_n(t) + B_i(t) + \tau_{in}(t) \tag{2.6}$$

$$A_n(t) = \nu_n(t) \tag{2.7}$$

$$B_i(t) = \sum_{k \in P_i} \gamma_k X_k(t) \beta_k + \tau_i(t) \tag{2.8}$$

$$\tau_{in}(t) \sim N(0, \theta_H) \qquad \nu_n(t) \sim N(0, \theta_A) \qquad \tau_i(t) \sim N(0, \theta_B). \tag{2.9}$$

2. Time Effect

$f(t)$ is a nonparametric function which depends only upon time. The covariate information is included through the $A_n(t)$ and $B_i(t)$ terms so the $f(t)$ describes the time effects not related to my time-varying covariates. In Bayesian adaptive regression splines [6], $f(t)$ is approximated by a cubic spline with k knots in locations $\xi = (\xi_1, \ldots, \xi_k)$ where $a < t_{(1)} < \xi_1 \leq \ldots \leq \xi_k < t_{(n)} < b$. Also, $b_j(t), j \in \{1, \ldots, k+2\}$ is the j^{th} function in a cubic B-spline basis with natural boundary constraints. Let $\boldsymbol{B}_{k,\xi}$ be the matrix whose i, j component is $b_j(t_i)$. Then $f(t) = \sum_{j=1}^{k+2} \omega_j b_j(t)$ for some $\omega_k, k \in \{1, \ldots, k+2\}$. The prior distributions are [7]:

$$p(k) = Poi(2) \tag{2.10}$$

$$p(\xi) = Unif(a, b) \tag{2.11}$$

$$p(\omega|k, \xi) = N\left(0, \theta_H n (\boldsymbol{B}_{k,\xi}^T \boldsymbol{B}_{k,\xi})^{-1}\right) \tag{2.12}$$

Using these choices for the prior distributions, ω can be integrated out of the posterior distribution to obtain a Markov chain for sampling from the marginal posterior of (k, ξ):

$$p(f(t)|k, \xi) = \int p(y|\omega, k, \xi) \pi(\omega|k, \xi) d\omega \tag{2.13}$$

The dimension of the posterior distribution of ξ is dependent upon k. To estimate the posterior, I use a reversible jump MCMC sampler [8, 9]. For each iteration of the sampler one of three moves is proposed: birth (add a new knot), death (remove an existing knot), or relocation (move an existing knot to a new location). The marginal posterior for (k, ξ) can be analytically computed. This makes it easy to compute the likelihood ratios $p(y|\xi^c, k^c)/p(y|\xi, k)$ which are used to determine if the sampler will move from state (ξ, k) to (ξ^c, k^c). For example, the likelihood ratio for the birth step is:

$$\frac{p(y|k^c, \xi^c)}{p(y|k, \xi)} = \frac{1}{\sqrt{n+1}} \left(\frac{y^T\{\boldsymbol{I}_n - n(n+1)^{-1}\boldsymbol{B}_{k,\xi}(\boldsymbol{B}_{k,\xi}^T\boldsymbol{B}_{k,\xi})^{-1}\boldsymbol{B}_{k,\xi}^T\}y}{y^T\{\boldsymbol{I}_n - n(n+1)^{-1}\boldsymbol{B}_{k^c,\xi^c}(\boldsymbol{B}_{k^c,\xi^c}^T\boldsymbol{B}_{k^c,\xi^c})^{-1}\boldsymbol{B}_{k^c,\xi^c}^T\}y} \right)^{n/2} \tag{2.14}$$

Additionally, the conditional posterior expectation is:

$$E[f(t)|k, \xi, y] = \frac{n}{n+1}\boldsymbol{B}_{k,\xi}(\boldsymbol{B}_{k,\xi}^T\boldsymbol{B}_{k,\xi})^{-1}\boldsymbol{B}_{k,\xi}^T y \tag{2.15}$$

This method performs well in my case, because the smoothness of the function is chosen automatically and not constrained to be constant across the domain. If there is a sharp change point in my data, this method will discover it. I greatly appreciate the software available from Robert Kass' website which greatly eased my implementation. For further information on the software and the implementation of this method, please see [10].

3. Determining the Significant Covariates

I have a large number of covariates (23), which are not free to obtain. I need to determine which covariates significantly contribute to the model. γ_k is a binary variable determining if covariate X_k is included in the model [11]. The prior and posterior distributions for γ are:

$$\frac{p(\gamma_k = 1|\gamma_{-k})}{p(\gamma_k = 0|\gamma_{-k})} = 1 \tag{2.16}$$

$$\frac{p(\gamma_k = 1|\gamma_{-k}, \beta, y)}{p(\gamma_k = 0|\gamma_{-k}, \beta, y)} = \frac{p(y|\gamma_k = 1, \gamma_{-k}, \beta)}{p(y|\gamma_k = 0, \gamma_{-k}, \beta)} \frac{p(\beta|\gamma_k = 1, \gamma_{-k})}{p(\beta|\gamma_k = 0, \gamma_{-k})} \tag{2.17}$$

where γ_{-k} is the rest of the γ-vector when the k^{th} element removed. When γ_k equals one, the covariate is included in the model. When it equals zero, it is excluded. Because I draw γ_k values from their posterior distribution in each iteration of the algorithm, the posterior probabilities of inclusion for each of the covariates are simply the proportion of draws which return a one.

It is likely that the set of significant covariates is dependent upon the order in which they are sampled. To overcome that potential problem, I randomly determine the order in which the γ_k values are sampled for each iteration and run multiple separate chain from disparate starting values to insure convergence.

4. Speed Parameter and Adoption Ceiling

The adoption ceiling is bounded both above and below. It is bounded above by 1 and below by the maximum cumulative adoption for the product-country pair observed in my data ($\max(Y_{in}(t))$). The prior distribution for α_{in} is taken to be uniform on that interval. The posterior distribution is proportional to:

$$p(\alpha_{in}|\cdot) \propto N\left[\frac{y_{in}(t)}{Y_{in}(t-1)} - \lambda_{in}(t) + \frac{\lambda_{in}(t)Y_{in}(t-1)}{\alpha_{in}M_i(t)}\middle| 0, \theta_L\right]. \tag{2.18}$$

The normalizing constant of the posterior distribution is not analytically tractable. I draw samples from that distribution using the Metropolis-Hastings sampler [12]. My proposal values are drawn independently from the prior distribution.

The speed parameter is constrained to the positive real line. The prior and posterior distributions are:

$$p(\lambda_{in}(t)) = Ga(\lambda_{in}(t)|0.001, 1000) \tag{2.19}$$

$$p(\lambda_{in}(t)|y) \propto N\left[\frac{y_{in}(t)}{Y_{in}(t-1)} - \lambda_{in}(t) + \frac{\lambda_{in}(t)Y_{in}(t-1)}{\alpha_{in}M_i(t)}\middle| 0, \theta_L\right]$$
$$\cdot N\left[\log\lambda_{in}(t) - f(t) - A_n(t) - B_i(t)|0, \theta_H\right]. \tag{2.20}$$

Again I draw samples from the posterior distribution using a Metropolis-Hastings sampler. The proposal values are obtained by adding a normal error to the current value on the log scale.

5. Precision Parameters

The precision parameters are given relatively noninformative prior distributions:

$$p(\theta_P) = Ga(\theta|10^{-5}, 10^{-5}) \qquad \text{for } P \in \{L, H, A, B\}. \tag{2.21}$$

The posterior distributions appear to be rather robust to the choice of hyperparameters. The full conditional posterior distributions are:

$$s_1^2 = \sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{t\in T_{in}} \log\left\{\frac{y_{in}(t)}{Y_{in}(t-1)\lambda_{in}(t)\left[1-\frac{Y_{in}(t-1)}{\alpha_{in}M_i(t)}\right]}\right\}^2 \tag{2.22}$$

$$p(\theta_L|\cdot) = Ga\left(\theta_L\left|10^{-5}+\frac{\sum_{n=1}^{N}\sum_{i=1}^{I}|T_{in}|}{2}, 10^{-5}+\frac{s_1^2}{2}\right.\right) \tag{2.23}$$

$$s_2^2 = \sum_{n=1}^{N}\sum_{i=1}^{I}\sum_{t\in T_{in}} \{\log[\lambda_{in}(t)] - f(t) - A_n(t) - B_{it}\}^2 \tag{2.24}$$

$$p(\theta_H|\cdot) = Ga\left(\theta_H\left|10^{-5}+\frac{\sum_{n=1}^{N}\sum_{i=1}^{I}|T_{in}|}{2}, 10^{-5}+\frac{s_2^2}{2}\right.\right) \tag{2.25}$$

$$p(\theta_A|\cdot) = Ga\left(\theta_A\left|10^{-5}+\frac{N}{2}, 10^{-5}+\frac{\sum_{n=1}^{N}\sum_{t\in T_n}A_n(t)^2}{2}\right.\right) \tag{2.26}$$

$$p(\theta_B|\cdot) = Ga\left(\theta_B\left|10^{-5}+\frac{I}{2}, 10^{-5}+\frac{\sum_{i=1}^{I}\sum_{t\in T_n}\left(B_{it}-\sum_{k\in P_i}\gamma_k X_k(t)\beta_k\right)^2}{2}\right.\right). \tag{2.27}$$

6. Random Effects

The $A_n(t)$ and $B_i(t)$ are the random country and product effects. The prior distributions are:

$$p(\beta) = N_k(\beta|0, \theta_B I) \tag{2.28}$$

$$p(A_n(t)) = N(A_n(t)|0, 10^{-5}) \tag{2.29}$$

$$p(B_i(t)) = N(B_i(t)|0, 10^{-5}). \tag{2.30}$$

The full conditionals are available and are:

$$p(\boldsymbol{\beta}|\cdot) = N_k\left(\boldsymbol{\beta}|(\boldsymbol{X}^T\boldsymbol{X}+\boldsymbol{I})^{-1}\boldsymbol{X}^T\boldsymbol{Y}, \theta_B(\boldsymbol{X}^T\boldsymbol{X}+\boldsymbol{I})\right) \tag{2.31}$$

$$p(A_n(t)|\cdot) = N\left(A_n(t)\left|\frac{N\theta_H\sum(\log\lambda_{in}(t)-f(t)-B_i(t))}{\theta_A+N\theta_H}, \theta_A+N\theta_H\right.\right) \tag{2.32}$$

$$p(B_i(t)|\cdot) = N\left(\mu_B, \theta_B+N\theta_H\right) \tag{2.33}$$

$$\mu_B = \frac{\theta_B\sum\gamma_k X_k(t)\beta_k + I\theta_H\sum(\log\lambda_{in}(t)-f(t)-A_n(t))}{\theta_B+N\theta_H}. \tag{2.34}$$

D. Results

1. Variable Selection Results

The variable selection results were surprisingly consistent between runs of the sampler. The only covariates with posterior inclusion probabilities greater than 0.5 are internet users, the consumer price index, and the pervasiveness of existing adopters. The three selected covariates all had inclusion probabilities greater than 0.9. The other covariates all had probabilities less than 0.15. The β values were also very consistent across sampler runs, with regular and unimodal posterior densities. The posterior means and standard deviations are summarized in Table IV. The number of internet

Table IV. Regression Coefficients

Covariate	Mean	Standard Deviation
Internet Users	0.199	0.008
Consumer Price Index	-0.168	0.006
Introductory Lag Year	0.332	0.034

users is positively related to the speed of the diffusion. The proliferation of the internet enables consumers to spread information more quickly and try or purchase products more easily. Additionally, one of the products I investigate is the personal computer which became much more useful with the internet.

The consumer price index is negatively related to the speed. When the cost of living is higher, consumers have less discretionary income to adopt new products.

Finally, when there are more adopters to make recommendations, the satisfaction and familiarity of the new product will increase [13]. Therefore the pervasiveness of existing adopters is positively related to the speed. Similar to common practice in the marketing literature [14], I operationalize the pervasiveness by the number of years the product introduction lags behind the global introduction in the lead country.

2. Time Component

I am mostly interested in the change in the speed over time. My hierarchical structure allows me to simply examine the estimate of $f(t)$ to see how the speed changes. There are two contrasting ways I could measure time, calendar year and year since introduction. Also, I could remove $f(t)$ from the model completely. I fit the model using each measure and then compared the models using DIC [15]. Table V describes the results of the model comparison. $\bar{D}$ is a measure of how well the model fits the

Table V. DIC Results

Time Measure	$\bar{D}$	P_D	DIC
None	33,280.6	600.3	33,880.9
Calendar Year	33,133.4	734.1	33,867.5
Year since Introduction	24,816.0	602.4	25,418.4

data. P_D is the effective number of parameters, which is used as a complexity penalty. DIC is the sum of those two values. In all cases, a smaller number is better.

Using calendar year provides a marginally better fit, but increases the effective number of parameters dramatically causing it to be preferred only slightly over the model without a time-varying component. The model using year since introduction vastly improves the fit of the model with only a nominal increase to the effective number of parameters. I use that model for the remainder of the chapter. Figure 1 summarizes the estimated posterior distribution of $f(t)$. The solid line is the point-

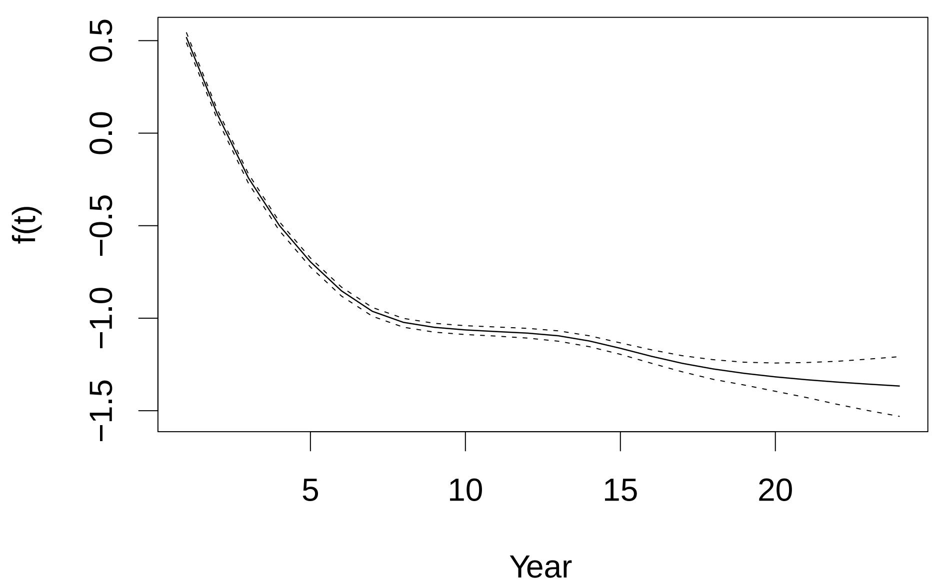

Fig. 1. $f(t)$ Against Year Since Introduction

wise posterior mean, and the dashed lines are the 95% credible interval bounds. The speed parameter starts high and then decreases until about year seven and becomes relatively constant. This effect makes sense because of initial promotion and buzz when a product is first introduced into a country.

The nonparametric function is not the only change to the speed parameter over time. Because two of the selected covariates (internet users and consumer price

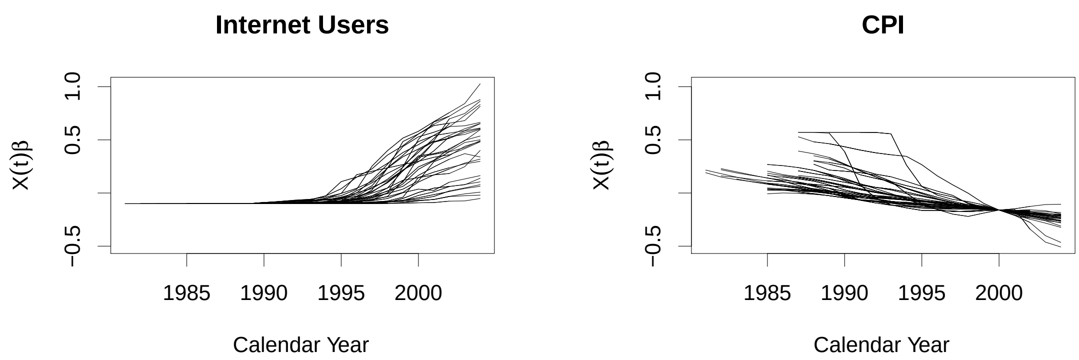

Fig. 2. Time-varying Regression Coefficients over Time

index) are both time-varying I need to include them when looking at the change of the speed over time. Figure 2 plots the two selected covariates against calendar year for each product and country pair. Please note that most the covariates (internet users included) have been standardized to have a mean of zero and a variance of one to improve the efficiency of the estimation. Consumer price index is calibrated by setting the year 2000 value to zero.

The number of internet users first grew above zero in 1989, but did not dramatically increase until the introduction of Netscape in 1995. Now that I have all the individual time-varying components, I can combine them (and include the lag effect covariate) to get a full picture of how the diffusion speed has changed over time (Figure 3). In the early years of each product introduction, the speed parameter is dominated by the $f(t)$ term, but as the product ages the internet effect pushes the speed parameter up.

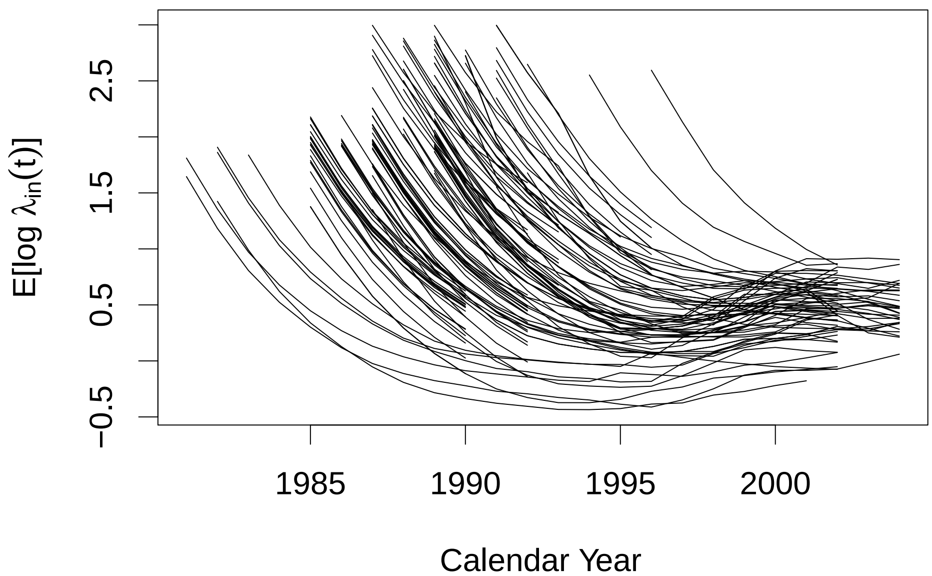

Fig. 3. Expected Time-varying Speed

E. Conclusion

Understanding the dynamic nature of new product diffusion speed is essential for global marketing managers to make informed decisions. [4] added substantially to that knowledge by showing how diffusion speed has increased in the United States. By relaxing his assumption of a constant speed parameter over the life of the product, I am able to show that the speed parameter is generally higher at introduction, falling to a low in the middle of the product's life and increasing again as technology has improved and our world has become more connected. Also, my global dataset allows us to show that this phenomenon occurs not just in developed nations (such as the

United States) but also in developing ones. Finally, I showed that the important covariates from my subset are the number of internet users, the consumer price index, and the number of years a country lags behind the global introduction of a product.

CHAPTER III

INVESTIGATING CROSS-COUNTRY INFLUENCE DYNAMICS

A. Introduction

There exists a rich stream of research on country-level new product diffusion [e. g. 16, 3, 17, 18, 2, 4]. At the same time, much of this research stream is focused on investigating the within-country diffusion process to understand the roles of macro environmental variables in driving differences in such diffusion across countries [19, 14, 20]. In contrast, as [21] observes, very little quantitative research has been conducted addressing the effect of cross-country interaction on the diffusion of new products. A similar observation is also made by [22] which emphasizes the need for additional research to better understand the spatial nature of the new product diffusion process, especially in the face of consistent empirical evidence on cross-country correlations in diffusion patterns. About a decade later, those observations still remain valid.

Many of the existing studies model cross-country influence in new product diffusion through product-specific word-of-mouth effects, either through learning/sequential models or the mixing/simultaneous models [23, 24, 14, 21, 13]. Also, those learning models assume a restrictive form of interaction based on the introduction lag period. They model the lead effect of early diffusion of a new product in one country (the lead country) on subsequent diffusion in another country (the lag country). Such an approach, while allowing for cross-country influence in analyzing diffusion, provides little insight into the underlying dynamics of cross-country influence. Further, this approach allows for only sequential and pair-wise, one directional cross-country influence, that of one lead country on one lag country at a time [14, 21].

The above limitations of the learning or sequential models to motivated a study

[21], which sought to expand the cross-country influence framework by pioneering the mixing models. Their proposed mixing model allows for multiple and simultaneous cross-country influence in analyzing the new product diffusion process. This model was conceptualized around a basic, population-based approach to capturing cross-country influence and it still remains prominent in the new product diffusion literature [25]. However, as [21] notes, their proposed mixing model was an initial foray in capturing the underlying dynamics of cross-country influence in new product diffusion, and it was meant to encourage future research exploring both complementary and competing explanations in greater detail.

In order to shed more in-depth insights into cross-country influence dynamics, [21] also points out the need for such future research on alternative approaches to modeling cross-country influence to test such approaches on a broader set of new products (than their four), countries (than their 10 EC nations) and parameter co-variates (than their two). Since that study two more studies [14, 26] have investigated multiple and simultaneous cross-country or cross-market influences in analyzing new product diffusion. However, both studies implicitly use a population based approach to capture cross-country interactions and allow only product-specific word-of-mouth from existing adopters to be the source of cross-influence.

The literature was later expanded through neighbor sets [27]. The n countries which have the most in common with the country of interest are included in the neighbor set. The commonality is instituted using geographical distance (the n closest countries), trade flow (the n countries with the most trade), and cultural similarity (the n countries whose normative distance in the four Hofstede cultural dimensions[28] is smallest). [29] also looks at the various cultural dimensions and their effect on the parameters in the Bass model.

As the above discussions imply, an important aspect of new product diffusion in need of additional research is the role of cross-country influence dynamics. Addressing the need for such research is also very time relevant as firms compete in an increasingly flat world marked by a substantial jump since the 1990s of the flow of goods, people and information across countries [1]. For instance, between 1995 and 2005, the volume of world trade and tourism increased by about 76% and 72% respectively. Over the same decade, the volume of international phone traffic more than doubled. My study builds on the limited set of existing studies on the role of cross-country influence and extends them in the following three important ways.

First, while [27] incorporated various measures of distance into their model, all the countries selected as neighbors are assumed to have an equal effect on the diffusion in the country of interest. Additionally, the countries not selected in the neighbor set have no effect on the country of interest. My models include the distance between the countries. This allows all countries to have a different level of effect a priori. My study develops and tests several models which capture interaction dynamics among countries. Specifically, I model cross-country influence through multiple conduits: bilateral flow of people (tourism), bilateral flow of goods/service (trade), normative cultural distance (Hofstede), and spatial proximity.

Second, the only source of cross-country influence on new product diffusion in the existing literature is the product-specific word-of-mouth effect from existing adopters. However, the source of cross-country information flow and influence on the consumption behavior of other potential adopters is unlikely to be so. For instance, it can be observational learning among reference leaders and followers as an interacting group [30]. My study applies an explicit reference leader-follower hierarchical structure among countries as another source of cross-country influence independent of the traditional product-specific word-of-mouth.

Third, a key reason for the limited existing research is the difficulty of collecting relevant data, especially in terms of information on bilateral interactions among countries [21]. In that respect, another contribution of my study is the collection and use of a novel data set. The scale and scope of this data set enables my study to provide in-depth and generalizable insights into the issues I investigate [3]. Drawn from several sources such as the International Monetary Fund, United Nations, World Bank and World Tourism Organization, the data set covers seven new product diffusions across 31 countries over the last three decades. It includes bilateral trade and tourism data among the 31 countries as well as information on a large number of macro-environmental covariates. The 31 countries cover essentially all the major developed and developing countries, accounting for about 80% of the global economic output and 60% of the global population.

As in many of the existing studies that investigate cross-country influence in new product diffusion, I use the well-known Bass Diffusion Model (BDM) as my core model. I then develop several augmented versions of it to address the role of cross-country influence dynamics in a more comprehensive and realistic manner. I estimate the proposed models using Hierarchical Bayesian techniques and compare their forecasting accuracy. Given the nature of my data, my estimation technique is particularly suited to make efficient use of relevant information across countries and products [2]. Taken together, the scope of my empirical data and proposed estimation methods enables my study to investigate several important but hitherto unexplored dynamics of cross-country influence on the new product diffusion process. My study makes significant contributions to an area that remains quite under-researched, especially in the context of accelerated globalization trends.

The rest of the chapter is organized as follows. In the next section, I discuss my proposed models. Section C discusses the data and section 2 presents my empirical estimation approach of the proposed models and the results from my empirical analyses. Finally, section D concludes with a summary discussion on the key insights from my study and on future research directions.

B. Conceptual Framework and Proposed Models

1. Conceptual Framework

It is natural to expect that a new product diffusion in a country will be influenced by both within-country or internal factors and cross-country or external factors. As noted earlier, the primary focus of my study is to investigate the rich dynamics of cross-country influence on the new product diffusion.

New product adoption behaviors by individuals in one geographic neighborhood have been observed to influence adoption behaviors of those living in surrounding neighborhoods [31]. Similarly, in the real world context, countries do not exist in isolation; rather they can be conceptualized to co-exist as neighbors or members in a global society or international community [32, 33]. In such a global society, when it comes to consumption and adoption trends, there exists an implicit reference hierarchy [34, 30]. One common driver of such a hierarchy has been readily observable status signals in the form of relative affluence and consumption spending levels across countries. Also, as members of the global society, the countries enjoy direct interactions with each other through multiple conduits or dimensions [34, 33]. Examples of such conduits of cross-country interactions may range from bilateral spatial proximity to bilateral flows of goods, people and investments. It is worth recognizing here that the recent accelerated trends in globalization and the notion of a flat world imply that spatial proximity will likely play less of a role in cross-country interactions than the flow of goods, people and investments. In such a global society, I conceptualize the role of cross-country influence dynamics on the new product diffusion process to depend on three primary factors.

The first factor is the source or nature of cross country influence in the new product diffusion process. I conceptualize such influence source to take the following two independent forms. In one form, the source of influence of country i on country j is the status of country i relative to country j in the reference hierarchy of the type discussed above. In the other form, the source of influence of country i on country j is specific to the new product being analyzed and is the current level of market penetration by the new product in country i. This source captures the likely cross-country influence on potential adopters in country j due to product-specific word-of-mouth by the people who have already adopted the product in country i. I should note that this is the only source of cross-country influence that is captured by the existing few studies that do investigate cross-country influence on the new product diffusion process [14, 21]. In contrast, my framework allows for both the forms to act as independent sources of cross-country influence.

The second factor is the bilateral interactions between the countries, which serve as conduits for the flow of cross-country influence. In my study, I consider the following four forms as the basis of such bilateral cross-country interactions: spatial proximity, bilateral flow of people, bilateral flow of goods and services, and cultural distance. So, my framework allows cross-country interactions to be decomposed in terms of their explicit bilateral components and offers four alternative ways to measure the intensity of such interactions. It is important to recognize that bilateral interactions between countries will vary in intensity, instead of being dichotomous in nature. In other words, even if two countries interact with the same set of countries, they can still differ in their levels of bilateral interactions across the countries and thus in the quality or composition of their neighborhood interactions based on what level with which neighbor.

Finally, another key factor driving cross-country influence dynamics in new product diffusion is the heterogeneity in relative responsiveness among countries to external influence. Such response heterogeneity is similar to empirically observed heterogeneity in response to within-country word-of-mouth influence [2]. As I discuss later, a number of country-specific covariates are expected to determine the relative responsiveness of a country to cross-country influence depending on the nature of influence. For example, all else being the same, a country in a higher position in the reference hierarchy will be less responsive to any cross-country influence in their adoption and consumption decisions.

2. Proposed Diffusion Model Extensions

The Bass Diffusion Model (BDM) has been widely used in the existing literature for investigating new product diffusion in general [3]. Following my earlier discussed conceptual framework, I propose several modified versions of the BDM to capture the role of cross-country influence dynamics in a more comprehensive and realistic manner. The BDM can be expressed as:

$$y_{in}(t) = [\alpha_{in} M_i(t) - Y_{in}(t-1)] \left[p_{in} + q_{in} \frac{Y_{in}(t-1)}{\alpha_{in} M_i(t)} \right] \exp\left[\epsilon_{in}(t)\right] \tag{3.1}$$

where $y_{in}(t)$ is the adoption sales for year t in country i for new product n, $Y_{in}(t)$ is the cumulative adoption sales, and $M_i(t)$ is the country population. The three parameters of the model are the market penetration potential (α_{in}), the coefficient of innovation or external influence (p_{in}), and the coefficient of imitation or internal influence (q_{in}). $\epsilon_{in}(t)$ is a zero-mean error term.

The model proposed by [27] (hereafter referred to as ABC) captures cross-country influence dynamics on the new product diffusion process in a broader way than the BDM. In addition to allowing for cross-country influence from implicit non-word-of-mouth sources through the coefficient p_{in} as in the BDM, model ABC allows for product-specific word-of-mouth as a source of cross-country influence:

$$\begin{aligned} y_{in}(t) = {} & [\alpha_{in} M_i(t) - Y_{in}(t-1)] \\ & \cdot \left[p_{in} + q_{in} \frac{Y_{in}(t-1)}{\alpha_{in} M_i(t)} + s_{in} \sum_{j \neq i} b_{ij} \frac{Y_{jn}(t)}{M_j(t)} \right] \exp\left[\epsilon_{in}(t)\right], \end{aligned} \tag{3.2}$$

$$a_{ij} = \frac{w_{ij}}{\sum_i \sum_{i \neq j} w_{ij}} \tag{3.3}$$

where w_{ij} denotes the observed level of a chosen basis of bilateral interaction between country i and country j, while a_{ij} denotes the relative level of that basis. As noted in my conceptual framework, I use four different empirical measures for w_{ij}: bilateral tourism flow, bilateral trade flow, the reciprocal of the normative cultural distance, and the reciprocal of the distance between the population centroids for country i and country j.

As evident from the above model structures and the coefficient q_{in}, the within-country or internal influence on the diffusion process of a new product is captured in the BDM and ABC through a product-specific word-of-mouth based social contagion process between adopters and non-adopters within the country. On the other hand, while they implicitly allow for cross-country or external influence from non-word-of-mouth sources through the coefficient p_{in}, they are silent on the specific nature or source of such influence as well as the particular basis of cross-country interactions that serve as conduits for such influence outside of the cross-country word-of-mouth. I implement an augmented version of both the ABC and the BDM which captures additional cross-country interaction dynamics through a reference hierarchy. To model the cross-country effects, I use insights from international consumption behavior studies that indicate such hierarchy is driven by readily observable status signals in the form of relative affluence and consumption spending levels across countries [34, 30].

Accordingly, I use time-averaged GNP of countries (L_j^*) to capture their relative affluence level. Specifically, each country's relative status is measured on a normalized scale (L_j) representing the country's relative affluence. An alternative measure using per capita consumption expenditure level across countries is found to be highly correlated with the GNP-based measure, as expected. The augmented version of the BDM (A-BDM):

$$y_{in}(t) = [\alpha_{in} M_i(t) - Y_{in}(t-1)] \left[p_{in} + q_{in} \frac{Y_{in}(t-1)}{\alpha_{in} M_i(t)} + r_{in} \sum_{j \neq i} a_{ij} L_j \right] \exp\left[\epsilon_{in}(t)\right], \tag{3.4}$$

$$a_{ij} = \frac{w_{ij}}{\sum_i \sum_{i \neq j} w_{ij}}, \tag{3.5}$$

$$L_j = \frac{L_j^*}{\sum_J L_j^*}. \tag{3.6}$$

Above, the model parameter r_{in} is similar to the external influence parameter p_{in}, and captures the responsiveness to cross-country influence.

The final proposed model (A-ABC) captures cross-country influence dynamics in the most expansive way by adding the reference hierarchy to the ABC:

$$\begin{aligned} y_{in}(t) &= [\alpha_{in} M_i(t) - Y_{in}(t-1)] \\ &\cdot \left[p_{in} + q_{in} \frac{Y_{in}(t-1)}{\alpha_{in} M_i(t)} + r_{in} \sum_{j \neq i} a_{ij} L_j + s_{in} \sum_{j \neq i} b_{ij} \frac{Y_{jn}(t)}{M_j(t)} \right] \exp\left[\epsilon_{in}(t)\right]. \end{aligned} \tag{3.7}$$

The measures a_{ij} and b_{ij} for relative levels of bilateral interactions remain as defined in the A-BDM and ABC respectively. Consistent with my conceptual framework, cross-country influence in my proposed models is decomposed in terms of bilateral interactions enjoyed by the focal country with the other countries and the corresponding levels of influence exerted by the other countries. Also, bilateral interaction of one country with another country in my proposed models differs not only in terms of the country being interacted with but also how and at what level.

3. Expected Covariates for the Proposed Model Parameters

As noted earlier, not only are there very few studies in the existing literature that investigate cross-country influence dynamics in the new product diffusion process, but these studies allow for a limited set of parameter covariates [14, 21, 25]. My study analyzes a much larger set of parameter covariates.

Further, as my study explicitly investigates the role of cross-country interaction dynamics on new product diffusion, it enables me to generate more reliable and general insights into the roles of such covariates on the diffusion process. For each of the parameters in my earlier proposed diffusion models, Table VI lists the covariates I analyze and their expected directional impacts. I next discuss the conceptual rationale behind such expected roles of the covariates for each of the model parameters.

a. Parameter for Penetration Potential (α)

Economic theories and empirical evidence from the existing diffusion studies imply that consumers who adopt a new product are those who have: (1) the ability to pay, (2) the willingness to pay, and (3) access to the product [18]. So the covariates likely to play a role on the magnitude of the country-product specific penetration potential parameter, α_{in}, are those that influence consumers' ability and willingness to pay for the product as well as their access to the product.

I use three covariates to reflect consumers' ability to pay. First, I use average national per capita income (adjusted for purchasing power parity). At the same time, average per capita income sheds no information on the distribution of such income across the population within a country, and this distribution can have a considerable effect on the new product diffusion [18]. As [2] argues, for a given level of

Table VI. Expected Effect of Country-specific Covariates

Parameter	Covariate	Expected Effect
α	Average Per Capita Income	Positive
	Elderly Population Proportion	Positive
	GINI Index	Negative
	Urban Population	Positive
	Trade	Positive
	Cell-Phone x Telephone Mainlines	Negative
	Cell-phone x Price Basket for Fixed Line	?
	Fax x Telephone Mainlines	Positive
	VCR x TV penetration rate	Positive
	Camcorder x TV penetration rate	Positive
p and r	Average Per Capita Income	Negative
	Individualism Index	Positive
	Uncertainty Avoidance Index	Negative
q and s	Internet penetration rate	Positive
	TV penetration rate	Positive
	GINI Index	Negative
	Female Labor Participation	Positive
	Individualism Index	Positive
	Uncertainty Avoidance Index	Negative
	Introductory Lag	Positive

average income, a country with a higher concentration in income has fewer consumers with adequate purchasing power to adopt a new product. I use the GINI Index as the measure of national income concentration. Since higher values of GINI Index indicate higher concentrations, it is expected to have a negative effect on penetration potential. Finally, I use the national demographic profile in terms of the elderly (65 years or more) proportion of the population to get a measure of disposable income. As the elderly typically have lower basic expenditures, higher elderly proportion of population will suggest higher disposable income for a given level of national per capita income.

As for consumers' willingness to pay for a new product, it will increase with expected incremental benefit offered by the new product relative to the product that currently serves that need [18]. Accordingly, if consumers have limited access to an existing product, they may be more willing to adopt a new product that is a substitute to the existing product. If a consumer already owns a complementary product that is needed to use the new product, she will have higher willingness to adopt it. Based on this rationale, I would expect that the fixed phone line penetration level will have a negative effect on cell phone penetration potential, but a positive effect on fax penetration potential. Also, TV penetration level will have positive effects on VCR and camcorder penetration potentials. I also expect price of land phone services to have a positive effect on cell phone penetration. On the other hand, if the price of fixed phone services is positively correlated with that of cell phone services within a country, then it will have a negative own price effect on cell phone penetration.

Finally, following [2], I use trade as a percentage of national GDP and urban population as a percentage of national population as two country level covariates affecting consumers' relative access to a new product. The rationale follows from the fact that higher trade fosters more open and competitive economy which in turn enhances product access through increased production and distribution efficiency [35]. Similarly, studies in urban economics show that urban areas are more likely to enjoy greater production and distribution efficiency from better infrastructure and economies of scale [36]. Therefore, I expect new product penetration to be higher in countries with higher levels of trade and urbanization.

b. Parameters for Non-Word-of-Mouth Based External Influence (p and r)

The parameters p and r represent the responsiveness of a country's new product adoption decision to influence sources not based on word-of-mouth influence from existing adopters in other countries. As discussed in my conceptual framework, one important source of such non-word-of-mouth external influence is the reference leader-follower hierarchy structure across countries with respect to consumption behavior. Also, international consumption behavior studies show that consumers in poorer countries are more likely to seek out information on the consumption behavior prevalent in richer countries [37]. So I would expect p and r to be negatively correlated with national per capita income.

I also expect that national cultural traits which reflect consumers' inclination to learn from other societies and cultural groups will have a positive effect on responsiveness to both external and internal influence sources [30]. Two well known measures of differences in cultural traits across countries that are particularly relevant in this context are the Individualism Index and Uncertainty Avoidance Index [28]. A country with a high Individualism Index reflects a cultural trait among its people where everyone is expected to interact beyond their familiar groups to look after themselves and their immediate families. On the other hand, a country with a high Uncertainty Avoidance Index reflects an insular cultural trait among its people whereby they are more intolerant of opinions different from what they are used to and more likely to believe that 'there can only be one Truth and I have it' [28]. The aforesaid discussion would suggest that the Individualism Index will have a positive relationship on p and r, while the Uncertainty Avoidance Index will have a negative relationship.

c. Parameters for Product-Specific Word-of-Mouth Influences (q and s)

The parameters q and s in the traditional BDM and my modified proposed models represent the responsiveness of a country's new product adoption decision to word-of-mouth influence from adopters within and outside the country. Thus, factors which facilitate the flow of word-of-mouth based information will positively affect the parameters q and s. These include the relative level of communication media in a country. I use two covariates in my analysis to capture the country specific level of communication media. One is the TV penetration level which represents the more traditional communication media and the other is the Internet penetration level representing the new interactive media [38].

I also use four other covariates which capture societal characteristics that are likely to facilitate the flow of word-of-mouth based information among its people within a country. One covariate used is the GINI Index to capture population heterogeneity in terms of income based on the rationale that personal interaction and communication are facilitated within homogeneous populations [13]. Another covariate is the proportion of females in a country's labor force. As women enter the labor force in greater numbers, they have greater opportunities to interact with men and other women with a consequent facilitation of greater social communication [2]. The other two of these four covariates are the cultural measures in terms of the Individualism Index and the Uncertainty Avoidance Index. As discussed earlier in terms of what these indices reflect, I would expect that the Individualism Index will have a positive relationship on parameters q and s, while the Uncertainty Avoidance Index will have a negative relationship.

Another factor that will positively affect the parameters q and s for a country is greater persuasiveness of word-of-mouth recommendations from the existing adopters [2]. The persuasiveness of recommendations will increase with the satisfaction and familiarity of existing adopters with the new product [13]. I use the number of years that the new product introduction in a country lags behind the introduction in the lead country as an operational measure for the relative level of satisfaction and familiarity of existing adopters for a new product [14].

C. Empirical Analysis and Results

1. Estimation Methodology

As discussed in §2, I use my proposed conceptual framework to develop two augmented models to better capture the role of cross-country influence dynamics on new product diffusion. In estimating these models, I use a hierarchical structure to borrow strength from other estimates in the same country or with the same product, and thus improve the Bayesian parameter estimates. I begin by allowing the parameters to vary over the real line. I apply an exponential transformation to the p_{in}, q_{in}, r_{in}, and s_{in} parameters to map them from the positive real line to the full real line. I also apply a logit transformation to the α_{in} parameter to move it from 0-1 to the full real line. I will denote the transformed variables with a star as follows:

$$\alpha^*_{in} = \log \frac{\alpha_{in}}{1-\alpha_{in}} \quad p^*_{in} = \log p_{in} \quad q^*_{in} = \log q_{in} \quad r^*_{in} = \log r_{in} \quad s^*_{in} = \log s_{in}. \quad (3.8)$$

I then apply the hierarchical structure to the transformed variables. I divide the transformed parameters into several parts; country- and product-specific portions and an interaction regression term. The regressors are related to the hypotheses in section 3. Using the standard BDM for illustration:

$$\begin{bmatrix} \alpha_{in}^* \\ p_{in}^* \\ q_{in}^* \end{bmatrix} = \begin{bmatrix} \alpha_i^* + \alpha_n^* \\ p_i^* + p_n^* \\ q_i^* + q_n^* \end{bmatrix} + \begin{bmatrix} \boldsymbol{X}_{\alpha in}^T \boldsymbol{\gamma}^{\alpha} \\ \boldsymbol{X}_{pin}^T \boldsymbol{\gamma}^{p} \\ \boldsymbol{X}_{qin}^T \boldsymbol{\gamma}^{q} \end{bmatrix} + \begin{bmatrix} \pi_{\alpha in} \\ \pi_{pin} \\ \pi_{qin} \end{bmatrix} \qquad \begin{bmatrix} \pi_{\alpha in} \\ \pi_{pin} \\ \pi_{qin} \end{bmatrix} \sim MVN(0, \Sigma_1) \tag{3.9}$$

With the country and product effects further decomposed.

$$\begin{bmatrix} \alpha_i^* \\ p_i^* \\ q_i^* \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}_{\alpha i}^T \boldsymbol{\beta}^{\alpha} \\ \boldsymbol{X}_{pi}^T \boldsymbol{\beta}^{p} \\ \boldsymbol{X}_{qi}^T \boldsymbol{\beta}^{q} \end{bmatrix} + \begin{bmatrix} \pi_{\alpha i} \\ \pi_{pi} \\ \pi_{qi} \end{bmatrix} \qquad \begin{bmatrix} \pi_{\alpha i} \\ \pi_{pi} \\ \pi_{qi} \end{bmatrix} \sim MVN(0, \Sigma_2) \tag{3.10}$$

$$\begin{bmatrix} \alpha_n^* \\ p_n^* \\ q_n^* \end{bmatrix} = \begin{bmatrix} \pi_{\alpha n} \\ \pi_{pn} \\ \pi_{qn} \end{bmatrix} \qquad \begin{bmatrix} \pi_{\alpha n} \\ \pi_{pn} \\ \pi_{qn} \end{bmatrix} \sim MVN(0, \Sigma_3) \tag{3.11}$$

The error term has a normal prior with a product-specific variance term, $\epsilon_{in}(t) \sim N(0, \sigma_n^2)$.

The majority of the parameters' posterior distributions are estimated using

Gibbs sampling [39] with non-informative yet conjugate priors.

$$
\begin{aligned}
s_n^2 &= \sum_{i=1}^{I}\sum_{t=1}^{T}\left(y_{in}(t)-\hat{y}_{in}(t)\right)^2 \\
p(\sigma_n^2|\cdot) &= InvGamma\left(\frac{TI}{2},\frac{s_n^2}{2}\right) && (3.12)\\
\boldsymbol{R}_\lambda &= \begin{bmatrix}\alpha_{in}^*\\ p_{in}^*\\ q_{in}^*\end{bmatrix} - \begin{bmatrix}\alpha_i^*+\alpha_n^*\\ p_i^*+p_n^*\\ q_i^*+q_n^*\end{bmatrix} - \begin{bmatrix}\boldsymbol{X}_{\alpha in}^T\boldsymbol{\gamma}^\alpha\\ \boldsymbol{X}_{pin}^T\boldsymbol{\gamma}^p\\ \boldsymbol{X}_{qin}^T\boldsymbol{\gamma}^q\end{bmatrix} \\
p(\lambda|\cdot) &= InvWish\left(5+NC, 0.1I_P+\sum_{n=1}^{N}\sum_{i=1}^{I}(\boldsymbol{R}_\lambda^T\boldsymbol{R}_\lambda)\right) && (3.13)\\
\boldsymbol{R}_\gamma &= \begin{bmatrix}\alpha_i^*\\ p_i^*\\ q_i^*\end{bmatrix} - \begin{bmatrix}\boldsymbol{X}_{\alpha i}^T\boldsymbol{\beta}^\alpha\\ \boldsymbol{X}_{pi}^T\boldsymbol{\beta}^p\\ \boldsymbol{X}_{qi}^T\boldsymbol{\beta}^q\end{bmatrix} \\
p(\gamma_i|\cdot) &= InvWish\left(5+I, 0.1\boldsymbol{I}_P+\sum_{i=1}^{I}(\boldsymbol{R}_\gamma^T\boldsymbol{R}_\gamma)\right) && (3.14)\\
p\left(\gamma_n|\cdot\right) &= InvWish\left(5+N, 0.1\boldsymbol{I}_P+\sum_{n=1}^{N}\begin{bmatrix}\alpha_n^*\\ p_n^*\\ q_n^*\end{bmatrix}^T\begin{bmatrix}\alpha_n^*\\ p_n^*\\ q_n^*\end{bmatrix}\right) && (3.15)\\
p(P_i^*|\cdot) &= N\big((\boldsymbol{\gamma}_i^{-1}+n\boldsymbol{\lambda}^{-1})^{-1}\cdot \\
&\qquad (\boldsymbol{\gamma}_i^{-1}\boldsymbol{X}_{Pi}^T\boldsymbol{\beta}^P+n\boldsymbol{\lambda}^{-1}(P_{in}^*-P_n^*-\boldsymbol{X}_{Pin}^T\boldsymbol{\gamma}^P)),\boldsymbol{\gamma}_i^{-1}+N\boldsymbol{\lambda}^{-1}\big) && (3.16)\\
p(P_n^*|\cdot) &= N\left((\boldsymbol{\gamma}_n^{-1}+n\boldsymbol{\lambda}^{-1})^{-1}(\boldsymbol{I}+\boldsymbol{\lambda}^{-1}(P_{in}^*-P_i^*-X_{Pin}^T\theta_P)),\boldsymbol{\gamma}_n^{-1}+n\boldsymbol{\lambda}^{-1}\right) && (3.17)\\
p(\beta_{iP}|\cdot) &= N\left((\boldsymbol{X}_{Pi}^T\boldsymbol{X}_{Pi})^{-1}\boldsymbol{X}_{Pi}^TP_i^*,(\boldsymbol{X}_{Pi}^T\boldsymbol{X}_{Pi})^{-1}\boldsymbol{\gamma}_i\right) && (3.18)
\end{aligned}
$$

where $P \in \{\alpha, p, q, r, s\}$).

The priors were tested for robustness. Estimation of the posterior distributions for α_{in}, p_{in}, q_{in}, r_{in}, and s_{in} require Metropolis-Hastings samplers[12], again with noninformative priors:

$$p(\alpha_{in}|\cdot) \propto \exp\Bigg\{ - \frac{\sum_{t=1}^{T}(y_{in}(t) - \hat{y}_{in}(t))^2}{2\sigma_n^2} - \frac{(\text{logit}(\alpha_{in}) - \alpha_i^* - \alpha_n^* - \boldsymbol{X}_{\alpha in}^T \boldsymbol{\gamma}^{\alpha})^2}{2\gamma_{1,1}} - \log(\alpha_{in} - \alpha_{in}^2)\Bigg\} \tag{3.19}$$

$$p(p_{in}|\cdot) \propto \exp\Bigg\{ - \frac{\sum_{t=1}^{T}(y_{in}(t) - \hat{y}_{in}(t))^2}{2\sigma_n^2} - \frac{(\log(p_{in}) - p_i^* - p_n^* - \boldsymbol{X}_{pin}^T \boldsymbol{\gamma}^{p})^2}{2\gamma_{2,2}} - \log(p_{in})\Bigg\} \tag{3.20}$$

$$p(q_{in}|\cdot) \propto \exp\Bigg\{ - \frac{\sum_{t=1}^{T}(y_{in}(t) - \hat{y}_{in}(t))^2}{2\sigma_n^2} - \frac{(\text{logit}(q_{in}) - q_i^* - q_n^* - \boldsymbol{X}_{qin}^T \boldsymbol{\gamma}^{q})^2}{2\gamma_{3,3}} - \log(q_{in})\Bigg\} \tag{3.21}$$

$$p(r_{in}|\cdot) \propto \exp\Bigg\{ - \frac{\sum_{t=1}^{T}(y_{in}(t) - \hat{y}_{in}(t))^2}{2\sigma_n^2} - \frac{(\text{logit}(r_{in}) - r_i^* - r_n^* - \boldsymbol{X}_{rin}^T \boldsymbol{\gamma}^{r})^2}{2\gamma_{4,4}} - \log(r_{in})\Bigg\} \tag{3.22}$$

$$p(s_{in}|\cdot) \propto \exp\Bigg\{ - \frac{\sum_{t=1}^{T}(y_{in}(t) - \hat{y}_{in}(t))^2}{2\sigma_n^2} - \frac{(\text{logit}(s_{in}) - s_i^* - s_n^* - \boldsymbol{X}_{sin}^T \boldsymbol{\gamma}^{s})^2}{2\gamma_{5,5}} - \log(s_{in})\Bigg\} \tag{3.23}$$

where the proposal values are generated through a random walk with variance determined in trial runs.

Using this method allows me to avoid many of the problems caused by the intractability of the likelihoods of models ABC and A-ABC. This occurs because I do not need to find the distribution of the parameters explicitly.

2. Model Estimation Results

a. Forecast Performance

Here I am interested in analyzing which of my four models best describes the underlying diffusion process. The BDM serves as a good benchmark model not only because my proposed models are extended versions of this popular diffusion model, but also for its use as the model of choice in [2].

Due to the hierarchical nature of my model framework, many forms of model selection (AIC, BIC, etc.) are difficult to compute because determining the number of parameters used is tricky. Other methods have been developed such as CPOs [40], Marginal Likelihood [41], and Reversible Jump MCMC [8]. While those methods are effective, there is a much simpler method which fits my situation well. I am able to compare the various models by how well they predict future values of the diffusion process. This makes more practical sense as well because a main purpose of this model is to be able to predict future values of the diffusion process. I predicted the diffusion level for one, two and three years beyond my sample, years 8-10 of the product life. The mean square prediction errors (MSPE) allow me compare the effectiveness of the various models and is calculated as follows:

$$\text{MSPE} = \left(\frac{y_{in}(t) - \hat{y}_{in}(t)}{M_i(t)}\right)^2 \tag{3.24}$$

where $\hat{y}_{in}(t)$ is the predicted value of $y_{in}(t)$. I then find the average MSPE over each country, product, year, and parameter draw. Table VII gives the MSPE multiplied by 10,000 for ease of comparison. Additionally, I calculated the average improvement over the base BDM and sorted the models by this improvement. I highlighted the lowest MSPE in each model for each prediction level.

For A-BDM, the centroid distance covariate performed the best, but not signifi-

Table VII. Model Comparison

Model	W-o-M	Reference	1 Year	2 Years	3 Years	Average	Improvement over BDM	Improvement over ABC
BDM			2.200	9.005	20.873	10.692		
A-BDM		Distance	0.931	3.714	**14.509**	**6.385**	40.29%	
		Tourism	**0.727**	**3.082**	15.928	6.579	38.47%	
		Cultural	0.852	3.712	20.042	8.202	23.29%	
		Trade	0.870	4.133	25.181	10.061	5.90%	
ABC	Cultural		**0.709**	**2.676**	**10.248**	**4.544**	57.50%	
	Distance		1.910	7.781	16.723	8.805	17.66%	
	Tourism		2.126	8.586	19.468	10.060	5.91%	
	Trade		2.137	8.676	19.914	10.242	4.21%	
A-ABC	Trade	Tourism	0.683	**2.329**	**6.815**	**3.275**	69.37%	27.92%
	Tourism	Distance	0.702	2.438	7.618	3.586	66.46%	21.09%
	Cultural	Distance	0.659	2.458	8.049	3.722	65.19%	18.10%
	Cultural	Cultural	**0.637**	2.432	8.766	3.945	63.10%	13.18%
	Cultural	Tourism	0.674	2.455	9.033	4.054	62.08%	10.78%
	Tourism	Trade	0.696	2.688	9.534	4.306	59.73%	5.24%
	Distance	Distance	0.654	2.449	10.146	4.416	58.70%	2.81%
	Trade	Distance	0.677	2.725	12.840	5.414	49.36%	
	Trade	Trade	0.716	2.953	14.172	5.947	44.38%	
	Tourism	Cultural	0.729	3.034	14.233	5.999	43.90%	
	Distance	Tourism	0.765	3.181	15.646	6.531	38.92%	
	Trade	Cultural	1.405	5.453	13.912	6.923	35.25%	
	Distance	Cultural	1.516	5.941	14.639	7.365	31.12%	
	Cultural	Trade	1.921	7.895	17.259	9.025	15.60%	
	Distance	Trade	2.208	8.854	19.945	10.336	3.34%	
	Tourism	Tourism	2.216	8.899	19.918	10.344	3.26%	

cantly more than the tourism covariate. In the first two years, the tourism covariate outperformed distance, but the distance measure was the best in year three. Both covariates outperformed the BDM by an average of 40%. In ABC, the cultural covariate vastly outperformed the other three options, outperforming the BDM by 57%. When looking at A-ABC, trade flow best described the word-of-mouth and tourism flow describes the reference hierarchy. While those two covariates performed suboptimally individually, there is an interaction effect which makes them the best when they are combined, outperforming the BDM by 69% and ABC by 28%. Notice also that the cultural and distance covariates performed relatively well in A-ABC, improving on ABC by 18%.

Because the MSPE is the smallest for A-ABC with trade and tourism as the covariates, I feel that the model represents the best description of the underlying diffusion process. This also makes heuristic sense. Tourism flow describes the reference hierarchy. The best way for people to determine what products are being used in the reference countries is by visiting them. The word-of-mouth is a little trickier. It does make sense that countries will show that they are using a certain product by exporting it, but why did it perform so poorly in ABC?

b. Hierarchical Regression Results

As noted earlier, my study uses a larger set of covariates than any of the other existing studies investigating cross-country influence on the new product diffusion process. Table VIII gives the results of the various covariates used in my hierarchical regression analysis of the model parameters. For each of my proposed models (A-BDM and A-ABC), I show the results for the version that performs the best in terms of prediction accuracy discussed above. Many of the covariates analyzed are found to be statistically significant (95% credible interval does not contain 0) and in almost all cases have the expected directional impacts on the respective model parameters. Since the directional impacts of the covariates on respective parameters are consistent across the various estimated models, I discuss below the results in terms of my most extensive or full model, A-ABC. In terms of the covariates for the penetration potential parameter (α), I find the following covariates to be significant – per capita income, elderly population ratio, GINI index, international trade and urbanization. It is interesting that the per capita income effect is negative when it was expected to be positive. International trade (as % of GDP), which is highly correlated with income, has a positive effect on penetration level as expected. It is possible that colinearity caused the unexpected effect in per capita income. After controlling for average per capita income level, the positive effect of elderly population ratio is consistent with the expectation that a higher value of such ratio reflects higher proportion of disposable income. The GINI index has a negative impact confirming the expectation that a more inequitable income distribution adversely affects new product penetration potential. Urbanization, with a negative effect, is the only covariate for which I find the directional effect to be contrary to initial expectations. In this context, it is relevant to note though that several major developing countries (e.g., China, India) with lower level of urbanization have in fact experienced a higher penetration level at comparative stages of the diffusion process for mobile phones [2].

As for covariates for the parameters (q) of internal product-specific word-of-mouth influence, introductory lag is found to have a significant positive effect. The result is consistent with my expectations as well as evidence from past studies [13]. I also find the GINI index to have a negative effect, which is in line with the expectation that the word-of-mouth based social contagion process will be less effective through a population with lower income homogeneity. Internet penetration rate is also found to have a negative effect. With more internet availability, people are less likely to

Table VIII. Hierarchical Regression Results

Covariate	BDM			A-BDM Ref. Hier. - Distance			ABC W. of M. - Cultural			A-ABC Ref. Hier. - Tourism W. of M. - Trade		
Potential Penetration (α)												
Intercept	-3.613	(0.000)	***	-100.540	(0.000)	***	-90.285	(0.000)	***	-26.174	(0.000)	***
Per Capita Income	-0.090	(0.000)	***	-0.009	(0.462)		-0.641	(0.448)		-0.792	(0.018)	*
Elderly Population Ratio	-2.808	(0.000)	***	217.611	(0.000)	***	206.739	(0.000)	***	90.315	(0.000)	***
GINI Index	-0.018	(0.000)	***	-0.384	(0.002)	**	-0.045	(0.490)		-0.283	(0.027)	*
Urbanization	-0.001	(0.415)		0.009	(0.437)		-0.161	(0.000)	***	-0.374	(0.000)	***
International Trade	0.004	(0.000)	***	0.159	(0.000)	***	0.071	(0.000)	***	0.042	(0.000)	***
Mainlines on Cell Phones	0.581	(0.000)	***	0.741	(0.000)	***	0.675	(0.000)	***	0.698	(0.000)	***
Price of Line on Cell Phones	-1.203	(0.000)	***	-6.606	(0.000)	***	-6.820	(0.000)	***	-4.650	(0.000)	***
Telephone Mainlines on Fax	-0.004	(0.008)	**	-0.002	(0.254)		0.000	(0.434)		-0.001	(0.348)	
TV Penetration on VCR	0.007	(0.000)	***	0.016	(0.024)	*	0.035	(0.000)	***	0.057	(0.012)	*
TV Penetration on Camcorder	0.003	(0.001)	***	0.050	(0.128)		0.024	(0.130)		0.084	(0.196)	
Internal Influence (p)												
Intercept	-6.893	(0.000)	***	-10.843	(0.000)	***	-11.524	(0.000)	***	-10.578	(0.000)	***
Per Capita Income	-0.380	(0.000)	***	-0.382	(0.001)	***	-0.355	(0.000)	***	-2.541	(0.000)	***
Individualism Index	0.038	(0.000)	***	0.036	(0.000)	***	0.046	(0.000)	***	0.035	(0.000)	***
Uncertainty Avoidance Index	-0.007	(0.000)	***	-0.004	(0.032)	*	0.003	(0.223)		-0.005	(0.122)	
Within-Country W-of-M (q)												
Intercept	-8.944	(0.000)	***	-12.498	(0.000)	***	-14.044	(0.000)	***	-16.113	(0.000)	***
Internet Penetration Rate	-0.001	(0.000)	***	0.001	(0.354)		0.004	(0.040)	*	-0.017	(0.000)	***
TV Penetration Rate	0.006	(0.000)	***	0.003	(0.490)		-0.011	(0.114)		0.000	(0.389)	
GINI Index	-0.072	(0.000)	***	-0.088	(0.003)	**	-0.105	(0.000)	***	-0.094	(0.000)	***
Female Labor Participation	0.018	(0.000)	***	0.014	(0.177)		0.037	(0.000)	***	0.033	(0.018)	*
Individualism Index	0.018	(0.000)	***	0.014	(0.068)		0.020	(0.055)		0.028	(0.019)	*
Uncertainty Avoidance Index	-0.002	(0.009)	**	0.003	(0.215)		0.014	(0.000)	***	0.016	(0.456)	
Introductory Lag	0.321	(0.000)	***	0.598	(0.000)	***	0.502	(0.000)	***	0.783	(0.000)	***

Table VIII. Continued

Covariate	BDM	A-BDM Ref. Hier. - Distance	ABC W. of M. - Cultural	A-ABC Ref. Hier. - Tourism W. of M. - Trade
Reference Hierarchy (r)				
Intercept		-9.435 (0.000) ***		-15.930 (0.000) ***
Per Capita Income		0.084 (0.334)		0.019 (0.401)
Individualism Index		-0.004 (0.209)		0.083 (0.000) ***
Uncertainty Avoidance Index		-0.008 (0.151)		-0.006 (0.265)
Cross-Country W-of-M (s)				
Intercept			-6.554 (0.045) *	-17.549 (0.000) ***
Internet Penetration Rate			-0.016 (0.000) ***	0.013 (0.000) ***
TV Penetration Rate			0.134 (0.000) ***	0.096 (0.000) ***
GINI Index			0.034 (0.066)	0.046 (0.145)
Female Labor Participation			-0.125 (0.000) ***	-0.003 (0.258)
Individualism Index			-0.022 (0.000) ***	-0.013 (0.303)
Uncertainty Avoidance Index			-0.096 (0.000) ***	-0.012 (0.288)
Introductory Lag			0.196 (0.000) ***	0.284 (0.134)

be influenced by those in their country. In contrast, with respect to the parameter (s) of external product-specific word-of-mouth influence, internet penetration level is found to have a strong positive effect. The TV penetration rate also has a significant positive influence. My study is the first to present systematic evidence that the emergence of the Internet and television has boosted the effect of word-of-mouth based cross-country influence on new product diffusion [3].

Per capita income and the individualism index were found to be significant when describing the non-word-of-mouth based internal parameter (p). Per capita income has a negative effect implying that people in poorer countries are more likely to be effected by non-word-of-mouth influences. Also, countries which score higher on the individualism index are more likely to be influenced in ways other than word-of-mouth. With respect to the non-word-of-mouth based external influence parameter (r) which captures the effects of reference leader-follower hierarchy across countries, only the individualism index had a significant effect. Adoption behaviors in countries that score higher on the individualism index are found to be more responsive to such cross-country non-word-of-mouth influences.

For my various diffusion models, I also note the estimated values of the various model parameters in Table IX. Please note that I used the logit transformation which constrained α to be between 0 and 1. My BDM parameter estimates are similar to those in the past studies [2].

c. Variance Decomposition of Heterogeneity

Variance decomposition allows me to allocate the total variance in my estimates to various parts of my model [42]. The hierarchical structure of my model allows me to be able to divide the variance into five categories; unobserved product effects, observed and unobserved country effects and observed and unobserved product and

Table IX. Parameter Estimates

	BDM	A-BDM	ABC	A-ABC
α	0.1550 (0.0015, 1.0000)	0.8417 (0.0191, 1.0000)	0.8872 (0.0546, 1.0000)	0.8074 (0.0212, 1.0000)
p	0.0088 (0.0002, 0.0331)	0.0020 (0.0000, 0.0238)	0.0019 (0.0000, 0.0226)	0.0016 (0.0000, 0.0205)
q	0.1639 (0.0048, 1.0786)	0.1061 (0.0000, 0.8642)	0.0967 (0.0001, 0.8192)	0.0629 (0.0000, 0.7630)
r		0.0005 (0.0000, 0.0030)		0.0003 (0.0000, 0.0026)
s			0.0019 (0.0000, 0.0202)	0.0117 (0.0000, 0.0943)

country interaction effects. Since A-ABC with trade and tourism was determined to be the best performing model, I performed the variance decomposition on its output. Table X shows the results.

Table X. Variance Decomposition

	Product Effects	Country Effects		Interactions		
	Unobserved	Observed	Unobserved	Observed	Unobserved	Total
α^*	0.12 (0.00%)	153.91 (2.09%)	0.05 (0.00%)	7224.06 (97.91%)	0.01 (0.00%)	7378.15
p*	1.80 (14.42%)	10.23 (81.97%)	0.13 (1.07%)	NA	0.32 (2.53%)	12.48
q*	1.05 (6.15%)	3.03 (17.72%)	0.08 (0.47%)	12.67 (74.07%)	0.27 (1.58%)	17.10
r*	0.21 (4.48%)	4.30 (93.89%)	0.04 (0.92%)	NA	0.03 (0.72%)	4.58
s*	0.17 (2.32%)	5.18 (72.60%)	0.08 (1.14%)	1.67 (23.35%)	0.04 (0.60%)	7.14

For α^*, almost all of the variance is captured in the observed country effects and the observed interactions between products and countries, mostly in the interactions. This implies that the adoption ceiling can be estimated rather well by the covariates I have specified. Notice also, that the total variance of α^* is rather large when compared to the other parameters. This is partly because α^* is found through a logit transformation, while the others use a log transformation. As α approaches 0 or 1, as was the case in a few instances, a small change in α will cause a large change in α^*.

In p^* and r^*, the observed interactions cell is not applicable because I did not specify any covariates there. Almost all of the variance is found in the observed country effects, showing that this coefficient can be well estimated by looking at other product launches in the same country. For q^* almost all of the variance is found in the interactions between the product and the country. This shows that you cannot just look at similar products in another country, or other products in the same country to estimate q. Rather much of the information comes from the introductory lag. As for s^*, it is mainly described by the observed country effects, with some help from the observed product effects. In summary, all of the parameters are rather well-described by the covariates I have chosen and therefore the estimates will be rather good. In comparing the results to [2], it seems that much of the unobserved idiosyncratic variance has been explained in my setup. That is likely due to the additional covariates in my study.

d. Time-varying Parameters

Implicit in my model is the fact that all my parameters are constant across time. A few adjustments to my fully augmented model (A-ABC) allow my parameters to vary across time.

$$\begin{aligned} y_{in}(t) &= [\alpha_{in}(t)M_i(t) - Y_{in}(t-1)] \\ &\cdot \left[p_{in}(t) + q_{in}(t)\frac{Y_{in}(t-1)}{\alpha_{in}(t)M_i(t)} + r_{in}(t)\sum_{j\neq i} a_{ij}L_j + s_{in}(t)\sum_{j\neq i} b_{ij}\frac{Y_{jn}(t)}{M_j(t)}\right] \exp\left[\epsilon_{in}(t)\right] \end{aligned} \tag{3.25}$$

$$\begin{bmatrix} \alpha^*_{in}(t) \\ p^*_{in}(t) \\ q^*_{in}(t) \\ r^*_{in}(t) \\ s^*_{in}(t) \end{bmatrix} = \begin{bmatrix} \alpha^*_i + \alpha^*_n + \alpha^*_t \\ p^*_i + p^*_n + p^*_t \\ q^*_i + q^*_n + q^*_t \\ r^*_i + r^*_n + r^*_t \\ s^*_i + s^*_n + s^*_t \end{bmatrix} + \begin{bmatrix} \boldsymbol{X}^T_{\alpha in}\boldsymbol{\gamma}^{\alpha} \\ \boldsymbol{X}^T_{pin}\boldsymbol{\gamma}^{p} \\ \boldsymbol{X}^T_{qin}\boldsymbol{\gamma}^{q} \\ \boldsymbol{X}^T_{rin}\boldsymbol{\gamma}^{r} \\ \boldsymbol{X}^T_{sin}\boldsymbol{\gamma}^{s} \end{bmatrix} + \begin{bmatrix} \pi_{\alpha in} \\ \pi_{pin} \\ \pi_{qin} \\ \pi_{rin} \\ \pi_{sin} \end{bmatrix}, \quad \begin{bmatrix} \pi_{\alpha in} \\ \pi_{pin} \\ \pi_{qin} \\ \pi_{rin} \\ \pi_{sin} \end{bmatrix} \sim MVN(0, \Sigma_1) \tag{3.26}$$

There are two different ways to measure time, through the standard calendar year or the number of years from the introduction of the product in the country (lag year). Analyzing the parameters through both metrics enables a deeper understanding of the dynamics. The time-varying components are estimated through cubic splines and the estimates and 95% credible intervals are depicted in Figure 4. If you are able to

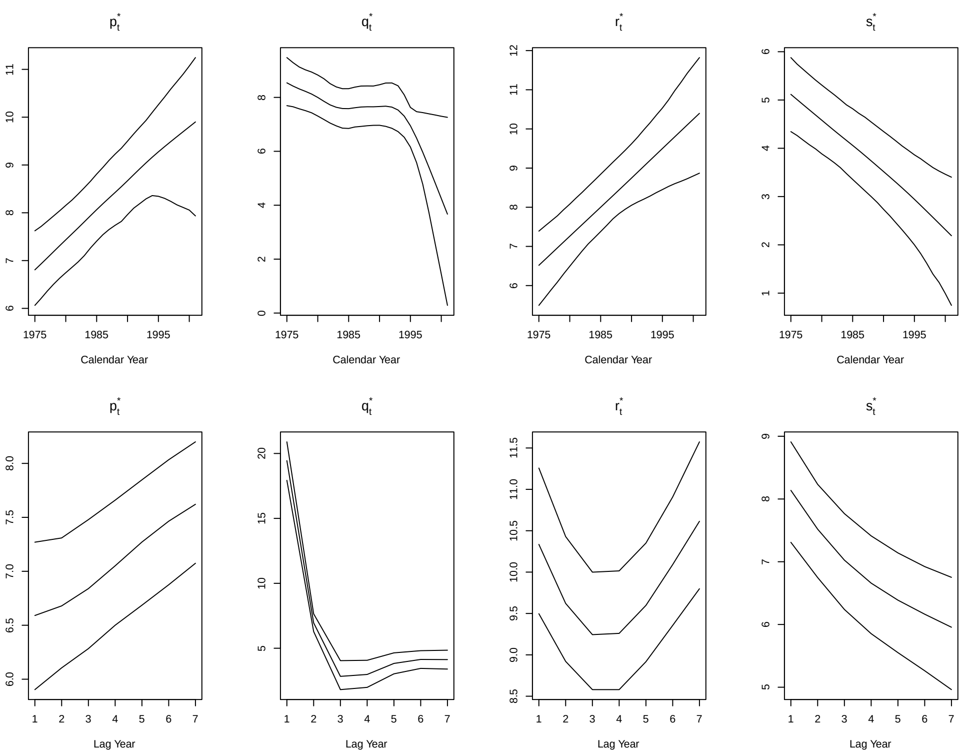

Fig. 4. Time-varying Parameters

draw a horizontal line across the plot and remain within the interval bounds, there is no evidence of a significant time effect. That is true for the s_t^* parameter under calendar years and for the p_t^*, r_t^*, s_t^* parameters under lag years. In both the p_t^* and r_t^* parameters, there is a distinct positive relationship over calendar time. With the increases in technology and globalization, it makes sense that the effect of non-word-of-mouth external influence would be increasing. The q_t^* parameter decreases over calendar time, which seems counter-intuitive. The real story is much clearer when the lag year effect is noticed. The parameter is highest at introduction and then quickly decreases. This is likely due to the buzz that is generated when a product is first introduced to a country. After the initial buzz dies off, the word-of-mouth effect is much smaller.

D. Conclusion

The new product diffusion process represents an important area of research in the marketing literature because of its obvious significance in understanding consumers' adoption behaviors and consequent strategic implications for firms. Not surprisingly, there exists a rich steam of marketing research on diffusion of new products [3, 43]. At the same time, this existing research remains quite limited in investigating the role of cross-country influence dynamics on the new product diffusion process. This limitation is particularly conspicuous in the context of recent acceleration in globalization trends and the surge in cross-country interactions in a flat world [1]. In other words, there remains an unfortunate disconnect between the widespread existence of cross-country interactions in reality and the very limited research in analyzing the expected role of such interactions on the new product diffusion process. The primary contribution of my study lies in taking an important and substantive step in addressing the aforesaid disconnect.

Consistent with the notion of simultaneous mixing models [21], my study proposes a conceptual framework to investigate the role of cross-country influence on the new product diffusion process in any country. Using the BDM as my core model, I apply this framework to develop 24 modified versions of the BDM by augmenting it with various competing and complementary forms – in both conceptual and empirical terms – of the role of cross-country influence dynamics. In developing my specific diffusion models from this framework, I use model structures that better capture the cross-country influence dynamics expected in reality. In that spirit, cross-country influence in my models is explicitly decomposed in terms of bilateral conduits of interactions enjoyed by a focal country and corresponding levels of influence exerted by other countries on the focal country. Specifically, I use models where such bilateral conduits of cross-country interactions are captured in terms of tourism flow, trade flow, cultural similarity, and spatial proximity between any two countries. Another distinctive aspect in my proposed models is the use of a reference leader-follower hierarchical structure among countries as an explicit source of cross-country influence on diffusion independent of the usual product-specific word-of-mouth.

I use hierarchical Bayesian techniques to estimate the models and then compare their relative predictive accuracies. The data used for my analysis are substantial in scope, covering seven new product diffusions across 31 countries with detailed information on bilateral trade and tourism flows across the countries. It also enables me to use the largest set of parameter covariates in investigating cross-country diffusion models to date. The sample set of countries covers essentially all the major developed and developing countries, and accounts for about 80% of the global economic output and 60% of the population.

My empirical analysis shows that almost all the proposed models that allow for explicit cross-country influence on new product diffusion consistently outperform the BDM in terms of relative predictive accuracy. This underscores the value of incorporating cross-country influence in diffusion models not only to better understand the dynamics of international diffusion, but also to improve the predictive power of such models. I find that the best performing model in terms of predictive accuracy is the one that allows for two sources of cross-country influence – the general reference leader-follower hierarchy among countries and the product-specific word-of-mouth from existing adopters. In terms of predictive accuracy, I also find that the best bilateral interaction conduits to capture the cross-country influence effects of reference leader-follower hierarchy and product-specific word-of-mouth are tourism and trade flows respectively. As for the results from the analysis of parameter covariates, I find that cultural and economic covariates are particularly significant in determining the responsiveness of countries to cross-country influences based on reference leader-follower hierarchy. I also find strong and systematic evidence that the emergence of the Internet has accentuated the effect of both within-country and cross-country influence of product-specific word-of-mouth from existing adopters. To my knowledge, my study is the first to document such evidence in new product diffusion studies [3].

My findings add new and substantive insights to the limited existing literature on the role of cross-country influence on the new product diffusion process. They are also of value to managers interested in better performing predictive models of international new product diffusion, especially in a world experiencing accelerated interactions among countries. I hope that my study stimulates additional studies on this under-researched but important issue of cross-country influence dynamics in the new product diffusion process.

In a broader context, my study is related to the understanding of social interactions and neighborhood influence dynamics in general, which has generated considerable interest among researchers in recent years with the emergence of social networking and digital communities [44]. While my study investigates such social interaction dynamics at macro level because of my focus on aggregate diffusion, an interesting area for future research would be to use micro level models to investigate the role of social interaction dynamics in a cross-country setting on individual consumers' new product adoption decisions. Another fruitful area for future research will be to analyze other forms of bilateral interactions like investment flows between countries in influencing aggregate diffusion, which can provide additional insights in developing an integrated theory of drivers for the new product diffusion process [3].

CHAPTER IV

A PRIOR DISTRIBUTION OF BAYESIAN NONPARAMETRICS INCORPORATING DISTANCES

A. Introduction

1. Problem Description

As mentioned in Chapter III, the Bass Diffusion model is commonly used in the new product diffusion literature. I restate the structure of the model in section B.

Implicit in the hierarchical formulation, all of the $\boldsymbol{\beta}$ parameters are assumed to be the same for each country. One additional dollar of GNPPC is assumed to have the same effect on the United States and the Philippines. I would like to allow the covariates to have a unique effect on each country. Fitting individual $\boldsymbol{\beta}$ values to each country would cause much of the information to be obtained from the prior, which in our case is meant to be relatively noninformative. I will group the countries into clusters using a Dirichlet process prior where the countries in the same cluster have the same $\boldsymbol{\beta}$ values but countries in different clusters have different $\boldsymbol{\beta}$ values. The organization of the countries will be updated in each iteration of the sampler creating a soft clustering. Alternatively, I could cluster the countries at the beginning of the analysis and fix the clusters while estimating the parameters. Soft clustering enables a nonparametric estimate of the $\boldsymbol{\beta}$ vector for each country, where the fixed clustering will only give an estimated vector for each cluster. Additionally, soft clustering incorporates the cluster uncertainty in the parameter estimates.

Under the standard Dirichlet process framework, each country is equally likely to be clustered with any other country. My method utilizes information about the countries to inform the prior clustering probabilities.

This new method will improve the out-of-sample prediction of the diffusion process.

2. Literature Review

There is a lot of research on the clustering of countries with respect to their macroeconomic variables [e.g. 45, 46, 47, 48, 49], the idea being that if you can find other countries that are similar, you can apply the same strategy and expect similar results. Alternatively, they focus their research efforts on a prototypical member of each group [50]. They often use either a single macroeconomic variable or use factor analysis or k-means clustering to aggregate the effect of multiple variables. There is little research using the actual diffusion process to cluster the countries, until [19].

This paper [19] has data from three consumer durables (TVs, VCRs, and CD Players) and 12 countries (Austria, Belgium, Denmark, France, Finland, the Netherlands, Norway, Sweden, Switzerland, the UK, Japan and the United States). For each product and country pair, they have sales data from introduction to 1990. They used 14 years of observations.

This [19] is the first paper which uses the diffusion process to cluster countries. They begin with the simple discrete time Bass model. First, they used factor analysis to compute the factor loadings of 23 macroeconomic variables. They used the mean values of the variables over time. The analysis divided the variables into five factors which were labeled as follows: overall mobility, health, trade, standard of living, and cosmopolitanism. Then they computed the factor scores for each of the 12 countries and then divided the scores via k-means clustering. They obtained a two-segment and three-segment solution. Both of the solutions are highly driven by the total mobility measurement. This part is similar to the previous work.

Next, they estimate the latent class model using the EM algorithm and AIC.

Using the EM Algorithm they estimate p_k, q_k, $\mathbf{\Sigma}_k$ and w_k, where $\mathbf{\Sigma}_k$ is the covariance matrix and w_k is the mixing proportion with likelihood defined as (4.1).

$$Y_c = \sum_{k=1}^{K} w_k f_{ck}(Y_c|\boldsymbol{X}_c, \boldsymbol{\beta}_k, \mathbf{\Sigma}_k) \tag{4.1}$$

Then they compute the posterior probability of country c being assigned to cluster k as (4.2).

$$P_{ck} = \frac{w_k f_{ck}(\cdot)}{\sum_{p=1}^{K} w_p f_{cp}(\cdot)} \tag{4.2}$$

Using the second method they find three clusters for color TVs, three for VCRs, and two for CD players. All of the posterior probabilities are at or close to one and zero. Finally, they measured the correspondence between the two clustering methods and found little connection. There is not much other work until 2009 [51].

The main focus of [51] is augmentation of the Bass model, but they do some clustering as an aside. Their data is a subset of 21 products across 70 countries for a total of 760 (52%) product and country combinations. Their method for computing the product diffusion curves includes some smoothing splines, which they call Augmented Functional Regression (4.3).

$$Y_i = \beta_0 + \sum_{j=1}^{4} g_j(e_{ij}) + \sum_{l=1}^{L-1} \delta_l I_{il} + \epsilon_{ij} \tag{4.3}$$

where the g_j are smoothing splines, the e_{ij} are the first four principal components scores, two each on the penetration curves $X_i(t)$ and on the velocity curves $X_i'(t)$, the I_{il} are product indicator variables, the ϵ_{ij} are normal errors and the δ_l are regression coefficients.

The majority of the paper is spent showing how this model outperforms many of the current models in use. But they do a little clustering by analyzing the first two e_{ij} and performing a k-means clustering. They get six clusters and then perform a few summary statistics comparing the clusters. They do not use the clustering to improve the fit of the model.

By incorporating a Dirichlet process prior for the $\boldsymbol{\beta}$ vector, I am able to cluster based on the diffusion process and use the clustering to improve the obtained predictions.

My work improves upon [19] by incorporating covariates into the model formulation. It improves upon [51] by using the clustering to improve my predictions. I improve upon both models through the use of model averaging and including the distance information in the prior distribution of the clusters.

B. Model

1. Standard Dirichlet Process

As mentioned in Chapter III, the Bass diffusion model [16] is commonly used in the new product diffusion literature. The hierarchical Bass diffusion model of [2] is:

$$y_{in}(t) = [\alpha_{in} M_i(t) - Y_{in}(t-1)] \left[p_{in} + q_{in} \frac{Y_{in}(t-1)}{\alpha_{in} M_i(t)} \right] \exp\left[\epsilon_{in}(t)\right] \tag{4.4}$$

where $y_{in}(t)$ is the adoption sales for year t in country i for new product n, $Y_{in}(t)$ is the cumulative adoption sales, and $M_i(t)$ is the country population. The three parameters of the model are the market penetration potential (α_{in}), the coefficient of innovation or external influence (p_{in}), and the coefficient of imitation or internal influence (q_{in}). $\epsilon_{in}(t)$ is a zero-mean error term. The model parameters are transposed to the real line:

$$\alpha^*_{in} = \operatorname{logit}(\alpha_{in}) \qquad p^*_{in} = \log(p_{in}) \qquad q^*_{in} = \log(q_{in}) \tag{4.5}$$

and the hierarchical structure defined as:

$$\begin{bmatrix} \alpha^*_{in} \\ p^*_{in} \\ q^*_{in} \end{bmatrix} = \begin{bmatrix} \alpha^*_i + \alpha^*_n \\ p^*_i + p^*_n \\ q^*_i + q^*_n \end{bmatrix} + \begin{bmatrix} \boldsymbol{X}^T_{\alpha in}\boldsymbol{\gamma}^{\alpha} \\ \boldsymbol{X}^T_{pin}\boldsymbol{\gamma}^{p} \\ \boldsymbol{X}^T_{qin}\boldsymbol{\gamma}^{q} \end{bmatrix} + \begin{bmatrix} \pi_{\alpha in} \\ \pi_{pin} \\ \pi_{qin} \end{bmatrix} \qquad \begin{bmatrix} \pi_{\alpha in} \\ \pi_{pin} \\ \pi_{qin} \end{bmatrix} \sim MVN(0, \Sigma_1) \tag{4.6}$$

With the country and product effects further decomposed.

$$\begin{bmatrix} \alpha^*_{i} \\ p^*_{i} \\ q^*_{i} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}^T_{\alpha i}\boldsymbol{\beta}^{\alpha} \\ \boldsymbol{X}^T_{pi}\boldsymbol{\beta}^{p} \\ \boldsymbol{X}^T_{qi}\boldsymbol{\beta}^{q} \end{bmatrix} + \begin{bmatrix} \pi_{\alpha i} \\ \pi_{pi} \\ \pi_{qi} \end{bmatrix} \qquad \begin{bmatrix} \pi_{\alpha i} \\ \pi_{pi} \\ \pi_{qi} \end{bmatrix} \sim MVN(0, \Sigma_2) \tag{4.7}$$

$$\begin{bmatrix} \alpha^*_{n} \\ p^*_{n} \\ q^*_{n} \end{bmatrix} = \begin{bmatrix} \pi_{\alpha n} \\ \pi_{pn} \\ \pi_{qn} \end{bmatrix} \qquad \begin{bmatrix} \pi_{\alpha n} \\ \pi_{pn} \\ \pi_{qn} \end{bmatrix} \sim MVN(0, \Sigma_3) \tag{4.8}$$

The covariates included in the model are the same as in Chapter III. Each country in a given cluster will have the same $\boldsymbol{\beta}^{\alpha}_i$, $\boldsymbol{\beta}^{p}_i$, and $\boldsymbol{\beta}^{q}_i$ so let $\boldsymbol{\beta}_i = (\boldsymbol{\beta}^{\alpha}_i, \boldsymbol{\beta}^{p}_i, \boldsymbol{\beta}^{q}_i)$. A common prior distribution for $\boldsymbol{\beta}_i$ has the form $N(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. I incorporate the Dirichlet process through the prior distribution:

$$\boldsymbol{\beta}_i | G \sim G \tag{4.9}$$

$$G \sim DP(\alpha_0 G_0) \tag{4.10}$$

$$G_0 = N_k(\boldsymbol{\theta}_0, \boldsymbol{\Sigma}_0). \tag{4.11}$$

With the addition of the Dirichlet process and the fact that I will now be clustering the countries into groups, I need to define more notation. Let $S_0 = \{1, 2, \ldots, n\}$ be the set of all the countries and a partition $\boldsymbol{\pi} = \{S_1, S_2, \ldots, S_l\}$ has the following properties:

1. $S_i \neq \emptyset$ for $i = 1, \ldots, l$ (all subsets non-empty)

2. $S_i \cap S_j = \emptyset$ for $i \neq j$ (mutually exclusive subsets)

3. $\cup_{j=1}^{l} S_j = S_0$ (exhaustive subsets)

$p(\pi)$ denotes a probability distribution for a random partition π. α_0 is the mass parameter which is related to the probability of an observation being assigned to a new cluster.

2. Incorporating a Single Distance

Under the standard Dirichlet process framework, each pair of countries have an equal chance of being clustered together *a priori*. I would like to incorporate the distances between each country into that prior probability.

The affinity distribution [52] rewards partitions which are formed by items which are close in terms of the distances d_{ij}. They define the affinity of items i and j as (4.12)

$$a_{ij} = \frac{(1+\epsilon)d^* - d_{ij}}{(1+\epsilon)d^*} \tag{4.12}$$

where d^* is the maximum distance among all pairs in the dataset. The kernel of the p.m.f. for π in the affinity distribution is:

$$p(\boldsymbol{\pi}) \propto \prod_{S \in \pi} \exp(t \cdot (g(S) - g(S_0))) \cdot [\alpha \Gamma(|S|)]^{1+t \cdot (g(S) - g(S_0))} \tag{4.13}$$

where $g(S)$ is the mean affinity of cluster S and $g(S_0)$ is the mean affinity of the cluster with all the observations together. t is the temperature which measures the magnitude of the distance effect. If the temperature were set to 0, the $p(\pi)$ would be equivalent to the formulation under the standard Dirichlet process.

3. Incorporating Multiple Distances

I can use all of the following as possible candidates in Chapter III to describe the distances:

- Trade between countries i and j

- Tourism flow between i and j
- Cultural Similarity
- Centroid Distance

The distance between country i and country j is then defined by the norm of a subset of the four distances:

$$d(i,j) = \sqrt{\sum_{k=1}^{K} \gamma_k d_k^2(i,j)} \tag{4.14}$$

where γ_k is a binary variable defining if the k^{th} covariate is included in the model and $d_k^2(i,j)$ is the square of the normalized distance in dimension k between countries i and j. Including γ_k allows me to be able to select among the four possible metrics. With only four metrics, I have $2^4 = 16$ possible γ vectors. I will fit the model with each of the possible vectors and then compare the predictive abilities to determine which subset is optimal.

4. Prior Specification

Many of the prior distributions remain the same as they were specified in Chapter III, but when incorporating a Dirichlet process, it is important that your prior is not overly vague. In each iteration of the Auxiliary Gibbs sampler [53], a parameter value for a currently unoccupied cluster is drawn from the prior distribution. If the prior is too vague, the values drawn will often be unreasonable and only rarely will a new cluster be formed.

In Chapter III, I use Zellner's g-prior. While that prior is not overly vague, it does require that the sample size is larger than the number of parameters. When all of the samples are in one cluster, that condition is satisfied. Once a single observation moves to a new cluster, the effective sample size becomes one and is smaller that the number of parameters.

For the joint prior of $(\boldsymbol{\beta}_i, \gamma_i)$, I use:

$$p(\boldsymbol{\beta}_i, \gamma_i) = Ng(\boldsymbol{\beta}_i, \gamma_i | 0, \boldsymbol{I}_k, \nu, \theta) = N_k(\boldsymbol{\beta}_i | 0, \gamma_i \boldsymbol{I}_k) Ga(\gamma_i | \nu, \theta). \tag{4.15}$$

The results were rather sensitive to the choice of hyperparameters in the gamma distribution. To obtain a draw from the prior distribution, I integrate out the γ_i value and obtain:

$$p(\boldsymbol{\beta}_i) = St_k \left(\boldsymbol{\beta}_i \middle| 0, \frac{\nu}{\theta} \boldsymbol{I}_k, 2\nu \right). \tag{4.16}$$

Analytically integrating parameters from the equations used in the sampler enables better mixing and faster (in the number of iterations) convergence. Although, computationally each iteration may take slightly longer.

The posterior distribution for the $\boldsymbol{\beta}$ values is also available with the γ_i integrated out:

$$\begin{aligned}
\boldsymbol{\theta}_n &= (\boldsymbol{I}_k + \boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{y} \\
\boldsymbol{\lambda}_n &= \boldsymbol{\theta} + \frac{1}{2} (\boldsymbol{y} - \boldsymbol{X} \boldsymbol{\theta}_n)^T \boldsymbol{y} \\
p(\boldsymbol{\beta}_i | \boldsymbol{X}, \boldsymbol{y}) &= St_k \left(\boldsymbol{\beta}_i \middle| \boldsymbol{\theta}_n, (\boldsymbol{I}_k + \boldsymbol{X}^T \boldsymbol{X}) \left(\nu + \frac{n}{2} \right) \boldsymbol{\lambda}_n^{-1}, 2\nu + n \right).
\end{aligned} \tag{4.17}$$

C. Results

1. Convergence Diagnostics

As with any MCMC method, it is important to ensure that the algorithm has converged so the obtained draws are coming from the posterior distribution. I measure convergence using entropy [54], defined as $-\sum_j (n_j/n) \log(n_j/n)$ where n_j is the number of observations in cluster j and n is the total number of observations. I run multiple chains, half commencing with all of the countries in one cluster and the other half with each country in its own cluster. All of the series seem to converged at around 40,000 iterations, so I will discard the first 40,000 iterations as burnin.

For each combination, I run multiple chains from various starting points to insure consistency of my estimates.

2. Temperature Comparison

I compare the average posterior number of clusters obtained for each of three temperature values in Table XI.

Table XI. Average Number of Posterior Clusters

Centroid	Cultural	Tourism	Trade	t=2	t=5	t=10
0	0	0	1	1.13	1.81	2.19
0	0	1	0	1.08	1.56	2.45
0	0	1	1	1.22	2.06	3.28
0	1	0	0	1.13	2.44	3.53
0	1	0	1	1.19	2.03	3.00
0	1	1	0	1.17	2.22	3.70
0	1	1	1	1.20	2.50	5.50
1	0	0	0	1.13	1.51	3.50
1	0	0	1	1.38	2.38	4.88
1	0	1	0	1.25	2.88	4.63
1	0	1	1	1.64	2.88	5.28
1	1	0	0	1.13	2.50	4.00
1	1	0	1	1.25	2.40	5.30
1	1	1	0	1.00	2.38	5.25
1	1	1	1	1.05	2.75	5.75

There is a definite increase in the number of clusters as the temperature increases. Additionally, as the temperature increases to 5 and again to 10, it seems there is some increase in the number of clusters as the number of distance measures increases.

3. Out-of-sample Prediction Results

The various distance measures are likely to be correlated. The sample Pearson correlations are depicted in Table XII. Because all of the measures are positively correlated,

Table XII. Distance Correlation Matrix

	Centroid	Cultural	Tourism	Trade
Centroid	1	0.165	0.369	0.335
Cultural	0.165	1	0.109	0.125
Tourism	0.369	0.109	1	0.629
Trade	0.335	0.125	0.629	1

it may be difficult to determine how much each distance measure contributes to the estimation.

To assess the utility of my method, I randomly remove one product-country data series (in this case, camcorders in Argentina) and then use the rest of the data to estimate the held-out portion. I will compare the mean absolute prediction error (MSPE) of the proportion of the population who have adopted the product for each possible subset of distance measures. When compared with the Bass diffusion model of [2] (with no clustering), the standard Dirichlet process improved the predictions by 2.2%. Table XIII describes how much each distance measure further improves the prediction for various temperature values. The distances are sorted by average performance, with the best performing subset first. The improvement values are in percentage improvement over the standard Dirichlet process with no distance information.

Incorporating the distance measures improved the vast majority of the predictions. The improvement is substantial, as much as 35%.

Table XIII. MSPE Improvement

Centroid	Cultural	Tourism	Trade	t=2	t=5	t=10
1	0	1	1	35.1%	14.8%	15.9%
0	1	0	0	22.4%	23.9%	18.9%
0	0	1	1	23.1%	16.9%	23.4%
0	1	0	0	29.0%	8.7%	21.3%
0	1	1	1	15.1%	27.8%	10.9%
0	1	1	0	17.5%	34.5%	1.3%
1	0	0	1	21.6%	23.2%	8.4%
0	0	1	0	1.0%	21.8%	22.2%
1	0	1	1	3.3%	17.1%	24.4%
0	0	0	1	11.8%	18.1%	3.5%
1	1	1	0	-1.7%	24.4%	4.5%
1	1	0	1	2.9%	16.5%	6.1%
1	0	0	0	27.3%	19.5%	-33.6%
1	1	1	1	-0.7%	25.3%	-16.7%
1	1	0	0	-25.2%	18.9%	13.5%

4. Posterior Clustering

To see how well the distance measures helped to cluster the countries, I look how they were clustered using the model with only the cultural distance measure included. Because I obtain a partition for each iteration of the sampler, I summarize the results using least squares clustering [55] in Table XIV.

Table XIV. Posterior Clustering

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Argentina	Australia	Canada	Chile	Denmark	Germany
Austria	United States	Netherlands	Portugal	Finland	Switzerland
Belgium				Norway	
Brazil				Sweden	
China					
France					
Greece					
Hong Kong					
India					
Ireland					
Italy					
Malaysia					
Mexico					
Philippines					
Singapore					
South Korea					
Spain					
Thailand					
United Kingdom					

It is interesting to note that in some clusters the cultural effect is very apparent (e.g. cluster 5 contains very similar countries), while others seem to have less in common (Chile and Portugal).

Having these clusters and the various distance metrics really help a manager when they are looking to enter a new country not in the current dataset. They are able to get probabilities that the country will be in each cluster. Using that information, they are able to get better predictions of the diffusion process in the new country.

D. Conclusion

Countries in today's global world are very diverse. Some developing nations would be greatly affected if five hundred more people had access to the internet. In more developed nations, that effect would be less pronounced. I have proposed a method to parsimoniously allow covariates to have different effects on disparate countries. My method improves the prediction of another product-country pair, allowing managers with established product lines in some countries to determine how their product would be adopted if they were to enter a new country.

CHAPTER V

CONCLUSION

Global marketing managers are keenly interested in being able to predict the sales of their new products. My work improves the understanding of how a product is adopted allowing the managers to optimally allocate their resources. My work explores how to describe the relationship between those countries and determines the best way to leverage that information to improve the sales predictions. I propose new marketing models in Chapters II and III which add flexibility without sacrificing interpretation. In Chapter IV, in addition to augmenting a current marketing model, I propose a new method for incorporating multiple distances into a nonparametric prior distribution. More precise descriptions follow.

In Chapter II, I that when compared to what is expected for the logistic diffusion model diffusion speed changes not only through the life of the product but also as a function of the calendar year. The speed is higher than expected at the introduction of the product, likely due to the initial promotion and buzz generated by an innovation. Additionally, as the internet has proliferated, the speed has increased.

In Chapter III, I show that by adding a reference hierarchy to the models of [2] and [27], the augmented models are better able to predict future sales of the product. Additionally, I show that the parameters in the Bass diffusion model change over time with the p and r parameters increasing over calendar time and the q and s parameters starting high at introduction and decreasing over the life of the product.

Finally, in Chapter IV, I describe a nonparametric method to cluster countries by their regression coefficients which improves the out-of-sample prediction.

These methods allow managers to better understand the new product diffusion process and make decisions which will maximize their stakeholders' value.

REFERENCES

[1] T. L. Friedman, *The World is Flat*. New York: Farrar, Straus and Giroux, 2005.

[2] D. Talukdar, K. Sudhir, and A. Ainslie, "Investigating new product diffusion across products and countries," *Marketing Science*, vol. 21, no. 1, pp. 97–114, 2002.

[3] D. Chandrasekaran and G. J. Tellis, "A critical review of marketing research on diffusion of new products," in *Review of Marketing Research*, N. K. Malhotra, Ed. New York: ME Sharpe, 2007, vol. 3, ch. 2, pp. 39–80.

[4] C. Van den Bulte, "New product diffusion acceleration: Measurement and analysis," *Marketing Science*, vol. 19, no. 4, pp. 366–380, 2000.

[5] W. James and C. Stein, "Estimation with quadratic loss," in *Proceedings of the Fourth Berkeley Symposium in Mathematical Statistics and Probability*, vol. 1, 1961, pp. 361–379.

[6] I. Dimatteo, C. R. Genovese, and R. E. Kass, "Bayesian curve-fitting with free-knot splines," *Biometrika*, vol. 88, no. 4, pp. 1055–1071, 2001.

[7] R. Kass and L. Wasserman, "A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion," *Journal of the American Statistical Association*, vol. 90, no. 431, pp. 928–934, 1995.

[8] P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, December 1995.

[9] D. G. T. Denison, B. K. Mallick, and A. F. M. Smith, "Automatic Bayesian curve fitting," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 60, no. 2, pp. 333–350, 1998.

[10] G. Wallstrom, J. Liebner, and R. E. Kass, "An implementation of Bayesian adaptive regression splines (BARS) in C with S and R wrappers," *Journal of Statistical Software*, vol. 26, no. 1, p. 1, 2008.

[11] P. Dellaportas, J. J. Forster, and I. Ntzoufras, "Bayesian variable selection using the Gibbs sampler," in *Generalized Linear Models: A Bayesian Perspective*, D. K. Dey, S. K. Ghosh, and B. K. Mallick, Eds. Marcel Dekker, Inc., 2000, pp. 273–286.

[12] W. K. Hastings, "Monte Carlo methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, April 1970.

[13] H. Takada and D. Jain, "Cross-national analysis of diffusion of consumer durable goods in Pacific Rim countries," *The Journal of Marketing*, vol. 55, no. 2, pp. 48–54, 1991.

[14] V. Kumar and T. V. Krishnan, "Multinational diffusion models: An alternative framework," *Marketing Science*, vol. 21, no. 3, pp. 318–330, 2002.

[15] D. J. Spiegelhalter, N. G. Best, B. P. Carlin, and A. Van der Linde, "Bayesian measures of model complexity and fit," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 64, no. 4, pp. 583–639, 2002.

[16] F. M. Bass, "A new product growth for model consumer durables," *Management Science*, vol. 15, no. 5, pp. 215–227, 1969.

[17] H. Gatignon, J. Eliashberg, and T. S. Robertson, “Modeling multinational diffusion patterns: An efficient methodology,” *Marketing Science*, vol. 8, no. 3, pp. 231–247, 1989.

[18] D. Horsky, “A diffusion model incorporating product benefits, price, income and information,” *Marketing Science*, vol. 9, no. 4, pp. 342–365, 1990.

[19] K. Helsen, K. Jedidi, and W. S. DeSarbo, “A new approach to country segmentation utilizing multinational diffusion patterns,” *The Journal of Marketing*, vol. 57, no. 4, pp. 60–71, 1993.

[20] G. J. Tellis, S. Stremersch, and E. Yin, “The international takeoff of new products: The role of economics, culture, and country innovativeness,” *Marketing Science*, vol. 22, no. 2, pp. 188–208, 2003.

[21] W. P. Putsis, S. Balasubramanian, E. H. Kaplan, and S. K. Sen, “Mixing behavior in cross-country diffusion,” *Marketing Science*, vol. 16, no. 4, pp. 354–369, 1997.

[22] M. G. Dekimpe, P. M. Parker, and M. Sarvary, “Multimarket and Global Diffusion,” in *New-product Diffusion Models*, ser. International Series in Quantitative Marketing, V. Mahajan, E. Muller, and Y. Wind, Eds. Boston: Kluwer Academic Pub, 2000, ch. 3, pp. 49–74.

[23] J. Ganesh and V. Kumar, “Capturing the cross-national learning effect: An analysis of an industrial technology diffusion,” *Journal of the Academy of Marketing Science*, vol. 24, no. 4, pp. 328–337, 1996.

[24] J. Ganesh, V. Kumar, and V. Subramaniam, “Learning effect in multinational

diffusion of consumer durables: An exploratory investigation," *Journal of the Academy of Marketing Science*, vol. 25, no. 3, pp. 214–228, 1997.

[25] Y. M. Van Everdingen, W. B. Aghina, and D. Fok, "Forecasting cross-population innovation diffusion: A Bayesian approach," *International Journal of Research in Marketing*, vol. 22, no. 3, pp. 293–308, 2005.

[26] C. Van den Bulte and Y. Joshi, "New product diffusion with influentials and imitators," *Marketing Science*, vol. 26, no. 3, pp. 400–421, 2007.

[27] P. Albuquerque, B. J. Bronnenberg, and C. J. Corbett, "A spatiotemporal analysis of the global diffusion of ISO 9000 and ISO 14000 certification," *Management Science*, vol. 53, no. 3, p. 451, 2007.

[28] G. Hofstede, *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Thousand Oaks CA: Sage Publications, 2001.

[29] C. Van den Bulte and S. Stremersch, "Social contagion and income heterogeneity in new product diffusion: A meta-analytic test," *Marketing Science*, vol. 23, no. 4, pp. 530–544, 2004.

[30] G. Ger and R. W. Belk, "I'd like to buy the world a coke: Consumptionscapes of the less affluent world," *Journal of Consumer Policy*, vol. 19, no. 3, pp. 271–304, 1996.

[31] A. Case, "Neighborhood influence and technological change," *Regional Science and Urban Economics*, vol. 22, no. 3, pp. 491–508, 1992.

[32] M. Featherstone, "Global culture: an introduction," *Theory, Culture & Society*, vol. 7, no. 2, pp. 1–14, 1990.

[33] R. Wilk, "Emulation, imitation, and global consumerism," *Organization & Environment*, vol. 11, no. 3, pp. 314–333, 1998.

[34] U. M. Dholakia and D. Talukdar, "How social influence affects consumption trends in emerging markets: An empirical investigation of the consumption convergence hypothesis," *Psychology and Marketing*, vol. 21, no. 10, pp. 775–797, 2004.

[35] I. W. Lieberman, "Privatization: The theme of the 1990s," *The Columbia Journal of World Business*, vol. 28, no. 1, pp. 8–17, 1993.

[36] P. S. Calem and G. A. Carlino, "Urban agglomeration economies in the presence of technical change," *Journal of Urban Economics*, vol. 29, no. 1, pp. 82–95, 1991.

[37] J. Tomlinson, *Cultural Imperialism: A Critical Introduction*. New York: Burns & Oates, 2002.

[38] B. T. Ratchford, D. Talukdar, and M. S. Lee, "The impact of the Internet on consumers' use of information sources for automobiles: A re-inquiry," *Journal of Consumer Research*, vol. 34, no. 1, pp. 111–119, 2007.

[39] A. E. Gelfand and A. F. M. Smith, "Sampling-based approaches to calculating marginal densities," *Journal of the American Statistical Association*, vol. 85, no. 410, pp. 398–409, 1990.

[40] A. E. Gelfand and D. K. Dey, "Bayesian model choice: Asymptotics and exact calculations," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 56, no. 3, pp. 501–514, 1994.

[41] S. Chib and I. Jeliazkov, "Marginal likelihood from the metropolis-hastings output," *Journal of the American Statistical Association*, vol. 96, no. 453, pp. 270–281, March 2001.

[42] J. E. Oakley and A. O'Hagan, "Probabilistic sensitivity analysis of complex models: A Bayesian approach," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 66, no. 3, pp. 751–769, 2004.

[43] V. Mahajan, E. Müller, and Y. Wind, Eds., *New-product Diffusion Models*, ser. International Series in Quantitative Marketing. Boston: Kluwer Academic Pub, 2000.

[44] W. R. Hartmann, P. Manchanda, H. Nair, M. Bothner, P. Dodds, D. Godes, K. Hosanagar, and C. Tucker, "Modeling social interactions: Identification, empirical methods and policy implications," *Marketing Letters*, vol. 19, no. 3, pp. 287–304, 2008.

[45] S. C. Jain, *International Marketing Management*, 4th ed. Boston: PWS-KENT, 1993.

[46] E. D. Jaffe, *Grouping: A Strategy for International Marketing.* New York: American Management Association, 1974.

[47] B. Toyne and P. G. P. Walters, *Global Marketing Management: A Strategic Perspective.* New York: Prentice Hall, 1989.

[48] A. C. Samli, R. Still, and J. S. Hill, *International Marketing.* New York: Macmillan, 1993.

[49] S. P. Sethi, "Comparative cluster analysis for world markets," *Journal of Marketing Research*, vol. 8, no. 3, pp. 348–354, 1971.

[50] J. Downham, "International market research," in *Consumer Market Research Handbook*, R. Worcester and J. Downham, Eds. New York: McGraw-Hill, 1986.

[51] A. Sood, G. M. James, and G. J. Tellis, "Functional regression: A new model for predicting market penetration of new products," *Marketing Science*, vol. 28, no. 1, pp. 36–51, 2009.

[52] D. B. Dahl, R. Day, and J. W. Tsai, "Distance-based probability distribution on set partitions with applications to protein structure prediction," paper presented at the Ninth Valencia International Meeting on Bayesian Statistics, Valencia, Spain, June 2010.

[53] R. M. Neal, "Markov chain sampling methods for Dirichlet process mixture models," *Journal of Computational and Graphical Statistics*, vol. 9, no. 2, pp. 249–265, Jun 2000.

[54] P. J. Green and S. Richardson, "Modelling heterogeneity with and without the Dirichlet process," *Scandinavian Journal of Statistics*, vol. 28, pp. 355–375, 2001.

[55] D. B. Dahl, "Model-based clustering for expression data via a Dirichlet process mixture model," in *Bayesian Inference for Gene Expression and Proteomics*, K. Do, P. Müller, and M. Vannucci, Eds. New York: Cambridge University Press, 2006.

VITA

Brian Matthew Hartman received his Bachelor of Science degree in actuarial science from Brigham Young University in 2006. He was named an Associate of the Society of Actuaries in 2008. He received his Doctor of Philosophy degree from Texas A&M University in 2010. His research interests include Bayesian Hierarchical Methods and Bayesian Nonparametrics.

He may be reached at Department of Statistics, 3143 TAMU, College Station, TX 77843-3143. His email address is brianmhartman@gmail.com